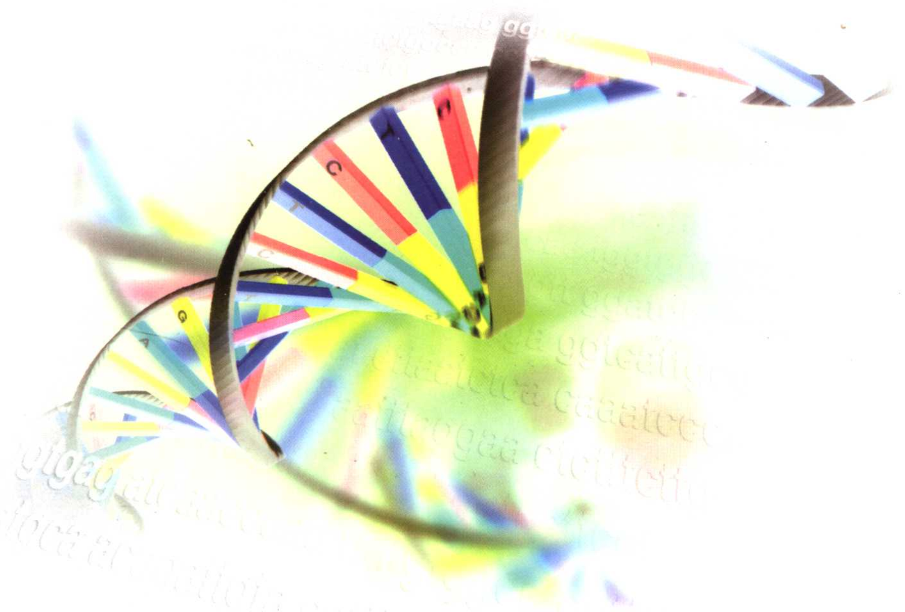
生物实验室系列

分子生物学实验技术

屈伸　刘志国　主编



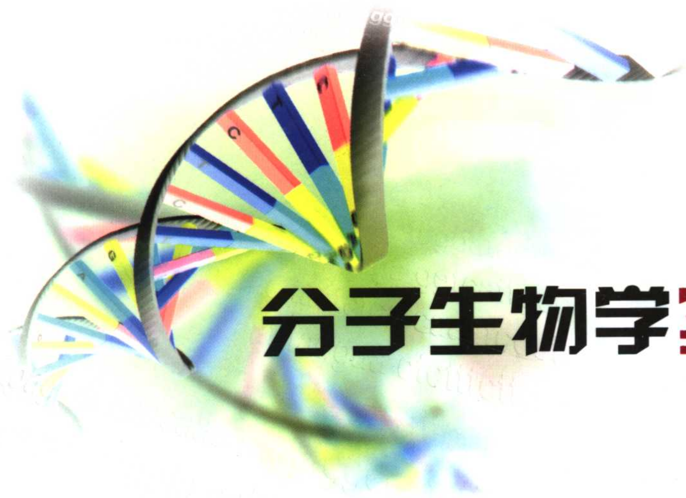
Chemical Industry Press

化学工业出版社
生物·医药出版分社

VISIT…

LANZAROTE Caliente.com

生物实验室系列



分子生物学实验技术

本书的三大特点

★ 本书将分子生物学实验相关技术完整地呈现在您面前，她能够助您完成从一名基本的实验技术入门者到一名深谙分子生物学实验技术奥妙的高级研究人员的转变，而且您也将发现学习分子生物学实验技术原来也可以这样有趣和轻松。

★ 对于初涉分子生物学领域的读者而言，她是一部必备的实验技术入门引导者，能够让您掌握基本的实验操作技术，如大分子制备和分析常用技术、蛋白质与核酸的提取与分离技术、分子标记技术等各项基本技术。

★ 对于欲探索分子生物学领域奥秘的科研工作者而言，本书也是一部得力的助手，她不仅详细阐述了各种技术的基本原理和操作，更着力于对相关理论基础进行深层次的剖析，以使融会贯通、举一反三，如生物信息学技术、分子改造技术、基因组学技术等。

销售分类建议：生物 / 生物技术
生物 / 生命科学

定价：65.00元

ISBN 978-7-122-01028-5

9 787122 010285 >

58.178
341

生物实验室系列

分子生物学实验技术

屈　伸　刘志国　主编

本书是关于分子生物学实验技术的一部综合性著作。它全面覆盖了有关分子生物学实验技术的诸多领域，包含了生物大分子制备和分析常用技术、蛋白质与核酸的提取与分离、PCR 技术、分子杂交与印迹技术、分子克隆技术、外源基因转移技术、蛋白质表达技术、分子标记技术、分子改造技术、测序及人工合成技术、基因组学技术、蛋白质组学技术、生物芯片技术、生物信息学技术、RNA 研究技术等。

作为一本实验技术类专著，本书不仅较详细地阐述了有关技术的具体操作和程序，更着力于对各种技术的基本原理及其相关理论基础进行深层次的剖析。故本书不仅可作为从事生命科学，特别是分子生物学相关领域工作者的实验室必备参考工具书，同时也可供高校相关专业师生及科学工作者对分子生物学实验技术的理论做深入探讨时参考。

图书在版编目（CIP）数据

分子生物学实验技术/屈伸，刘志国主编．—北京：化学工业出版社，2007.9

（生物实验室系列）

ISBN 978-7-122-01028-5

Ⅰ. 分…　Ⅱ. ①屈…②刘…　Ⅲ. 分子生物学-实验Ⅳ. Q7-33

中国版本图书馆 CIP 数据核字（2007）第 135059 号

责任编辑：郎红旗　周　旭　邵桂林

责任校对：宋　玮　　　　装帧设计：关　飞

出版发行：化学工业出版社　生物·医药出版分社（北京市东城区青年湖南街 13 号　邮政编码 100011）

印　　刷：大厂聚鑫印刷有限责任公司

装　　订：三河市延风装订厂

787mm×1092mm　1/16　印张 29¾　字数 722 千字　2008 年 1 月北京第 1 版第 1 次印刷

购书咨询：010-64518888（传真：010-64519686）　售后服务：010-64518899

网　　址：http://www.cip.com.cn

凡购买本书，如有缺损质量问题，本社销售中心负责调换。

定　　价：65.00 元

版权所有　违者必究

出版者的话

21 世纪是生命科学的世纪，这已成为人们的共识。

生命科学随着人类对自身和自然的认识、探索而萌芽，随着人类生产和科学实践的进步而发展。现代生命科学包括生物学、医学、农学等传统学科领域，以及生物学、生物技术与环境科学乃至社会科学等其他学科相互渗透、交叉而产生的新型学科体系。20 世纪后叶现代生物科学尤其是分子生物学取得了一系列突破性成就，使得生命科学在自然科学体系中的位置发生了革命性的变化，成为 21 世纪的带头学科。人们对生命科学也寄予了无限的期望，希望能够解决人类社会所面临的人口膨胀、资源匮乏、疾病危害、环境污染和生态破坏等一系列重大问题。

回顾生命科学的发展历程，实验技术一直起着非常重要的促进作用。如 17 世纪 Leeuwenhoek 等人发明并应用显微镜技术，直接催生了“细胞学说”的建立和发展；1973 年 Cohn 和 Boyer 完成了 DNA 体外重组实验，标志着基因工程的肇始；1988 年 Kary Mullis 发明的 PCR 技术甚至使生命科学产生了飞跃性的发展。可以说，生命科学无时无刻离不开实验，实验是开启神奇的生命王国大门的钥匙。没有实验技术的不断进步，也就没有生命科学今天的巨大发展；同时，生命科学的发展又对实验技术提出了更高的要求，进一步刺激了后者的不断进步。生命科学正是在“实验催生和验证着基础理论，理论指导和发展了实验技术”的不断循环中从必然王国走向自由王国。

工欲善其事，必先利其器。为了有助于生命科学工作者更多地了解相关实验技术和仪器设备，更好地设计实验方案，更有效地开展实验过程，更合理地处理实验结果，化学工业出版社组织出版了“生物实验室系列图书”。系列图书在整体规划的基础上，本着“经典、前沿、实用，理论与技术并重”的原则组织编写，分批出版。

在题材上，系列图书涵盖综合实验技术和单项实验技术两个方面。其中综合实验技术既有以实验目的为题，如“蛋白质化学分析技术”，内容纵向覆盖多项实验技术；也有以某一生命学科领域的综合实验技术为题，如“发酵工程实验技术”、“生物化学实验技术”等。而单项实验技术则以深入介绍某一专项技术及其应用为主，在阐述其基本原理的基础上，横向介绍该项技术在多个领域的应用，如“双向电泳技术”、“流式细胞术”等。

在内容上，系列图书主要有以下两个显著特点。一是强调先进性——除了系统介绍常用和经典实验技术以外，特别突出了当前该领域实验手段的新理论、新技术、新发展，为国内专业人员起到借鉴和引导作用。二是强调可操作性——对于每一项实验技术，系统介绍其原理方法、设备仪器和实验过程，让读者明了实验的目的、方案设计以及具体步骤和结果处理，以期起到实验指南的作用。

本系列图书坚持质量为先，开拓国内和国际两个出版资源。一方面，约请国内相关领域兼具理论造诣和丰富实验室工作经验的专家学者编著；另一方面，时刻关注国际生命科学前沿领域和先进技术的进展，及时引进（翻译或影印）国外知名出版社的权威力作。

“生物实验室系列图书”的读者对象设定为国内从事生命科学及生物技术和相关领域（如医学、药学、农学）的专业研究人员，企业或公司的生产、研发、管理技术人员，以及高校相关专业的教师、研究生等。

我们殷切希望“生物实验室系列图书”的出版能够服务于我国生命科学的发展需要，同时热忱欢迎从事和关心生命科学的广大科技人员不仅对已出版图书提供宝贵意见和建议，也能对系列图书的后续题目设计贡献良策或推荐作者，以便我们能够集思广益，将这一系列图书沿着可持续发展的方向不断丰富品种，推陈出新。

谨向所有关心和热爱生命科学，为生命科学的发展孜孜以求的科学工作者致以崇高的敬意！

祝愿我国的科技事业如生命之树根深叶茂，欣欣向荣！

化学工业出版社

生物·医药出版分社

前　言

1953 年 Watson 和 Crick 等提出 DNA 的双螺旋结构模型，阐明了遗传信息的传递规律，从而开辟了分子生物学的新纪元。20 世纪 70 年代中期出现的分子克隆技术进一步推动了分子生物学的快速发展。进入 21 世纪，随着人类基因组计划的顺利完成，分子生物学研究发展到基因组及后基因组时代，各种组学技术的发展加快了生物学领域中科学研究的进程，同时也带动和促进了许多新技术、新方法的发展，并与许多相关学科交叉、融合，派生出许多新的学科，如：分子遗传学、分子医学等。不仅如此，在应用领域，利用分子生物学原理和方法不仅可以大规模生产多种基因工程药物、基因工程疫苗，而且可以治疗人类遗传疾病，改良动、植物品种，及获得转基因动物和转基因植物等，创造出巨大的社会与经济价值。在新的世纪里，随着分子生物学研究的不断深入及新的研究方法和技术的大量涌现，分子生物学正成为带动生命科学发展、促进科技进步的重要学科力量。

本书较为系统地介绍了分子生物学领域中大量使用的实验技术原理和方法。从较为基础的生物分子的提取、制备及检测，到步骤相对复杂的分子克隆技术；从离心、电泳等通用技术，到针对性强的 PCR 技术、分子杂交等技术；从局部的、单个对象的研究，到群体和组学的研究等。这使本书成为一本较为完整的分子生物学实验技术和方法的参考书。在内容安排上，本书编写遵循分子生物学研究的一般顺序，从常用的基本的技术开始，如生物分子的提取和分析，过渡到较为独立的专门方法，如：PCR 技术、分子标记技术、分子杂交技术等，再到较为复杂的克隆技术，最后介绍最新的基因组学技术、蛋白质组学技术、生物信息学技术等。本书重点突出分子生物学实验技术的系统性和对实验的指导性，希望能为分子生物学及相关专业的科研、教学、技术人员和研究生提供一本较为系统全面的实验指南。

本书编写人员均从事分子生物学研究工作多年，具有丰富的分子生物学理论知识和研究经验。编写中注意做到理论与实践结合，注重突出重点和对实验的指导性。本书编写过程中得到了有关单位的大力支持，华中科技大学同济医学院、武汉大学病毒研究所、武汉工业学院等院校的教师为本书的编写提供了大量资料，化学工业出版社为本书的顺利出版做了大量的工作，在此，对上述单位及相关人员的大力支持表示衷心的感谢。由于分子生物学技术和方法的复杂性并处于快速发展中，以及编写人员知识的局限，书中难免出现缺点和错误，恳请读者提出宝贵意见。

编者

2007 年 6 月

目　录

第一章　生物大分子制备和分析常用技术 …… 1

第一节　概论 …… 1

一、预处理和细胞的分离 …… 1

（一）选择材料及预处理 …… 1

（二）细胞的分离 …… 2

二、细胞的破碎及细胞器的分离 …… 3

第二节　电泳技术 …… 5

一、基本原理 …… 6

二、影响电泳的因素 …… 7

三、醋酸纤维素薄膜电泳 …… 8

四、琼脂糖凝胶电泳 …… 9

（一）核酸分子大小与琼脂糖浓度的关系 …… 9

（二）核酸构型与琼脂糖凝胶电泳分离的关系 …… 9

（三）琼脂糖凝胶电泳基本方法 …… 10

五、常规聚丙烯酰胺凝胶电泳 …… 10

（一）聚丙烯酰胺凝胶聚合原理及相关特性 …… 11

（二）聚丙烯酰胺凝胶电泳原理 …… 12

（三）操作方法 …… 13

六、SDS-聚丙烯酰胺凝胶电泳原理 …… 14

七、聚丙烯酰胺梯度凝胶电泳 …… 15

八、聚丙烯酰胺凝胶等电聚焦电泳 …… 16

（一）等电聚焦的原理 …… 17

（二）pH 梯度的形成 …… 17

九、双向凝胶电泳 …… 18

十、染色方法 …… 19

（一）蛋白质染色 …… 19

（二）核酸染色 …… 20

第三节　色谱技术 …… 21

一、色谱技术的概念 …… 21

二、色谱法的分类 …… 21

（一）按两相所处的状态分类 …… 21

（二）按色谱的分离机制分类 …… 22

三、常用的色谱方法 …… 23

（一）凝胶色谱 …… 23

（二）离子交换色谱 …… 24

（三）亲和色谱 …… 26

（四）高效液相色谱 …… 27
（五）毛细管电泳 …… 28
第四节 离心技术 …… 31
一、离心理论 …… 32
（一）离心分离的原理 …… 32
（二）相对离心力和离心时间 …… 33
（三）离心机的分类 …… 35
二、离心分离的种类 …… 35
（一）差速离心法 …… 36
（二）密度梯度离心法 …… 36
三、离心操作注意事项 …… 37
第五节 分光光度技术 …… 38
一、基本原理——朗伯-比尔定律 …… 38
二、分光光度技术的应用 …… 39
（一）定量分析 …… 39
（二）定性分析 …… 40
第二章 蛋白质、核酸的提取与分离 …… 42
第一节 蛋白质的提取、分离纯化及定量 …… 42
一、蛋白质（包括酶）的提取 …… 42
（一）水溶液提取法 …… 42
（二）有机溶剂提取法 …… 43
二、蛋白质的分离纯化 …… 43
（一）根据蛋白质溶解度不同进行分离的方法 …… 44
（二）利用色谱技术对蛋白质进行分离纯化 …… 45
（三）利用电泳技术对蛋白质进行纯化分离 …… 48
（四）利用超速离心分离制备蛋白质 …… 48
（五）膜分离技术在蛋白质分离纯化中的应用 …… 49
（六）重组蛋白的分离纯化 …… 50
（七）分离制备蛋白质的一个实例 …… 52
三、蛋白质的定量 …… 55
第二节 DNA 的提取与分离 …… 57
一、质粒 DNA 的提取与分离 …… 57
（一）质粒提取的一般步骤 …… 57
（二）质粒 DNA 的小量制备 …… 58
（三）质粒 DNA 的大量制备 …… 59
二、染色体 DNA 的提取与分离 …… 61
（一）从培养细胞中提取 DNA …… 61
（二）从组织中提取 DNA …… 62
（三）从大量血样的白细胞中分离高分子量 DNA …… 62
（四）从小量血样的白细胞中分离高分子量 DNA …… 63
（五）高分子量 DNA 的小量快速制备 …… 63

第三节 RNA 的提取与分离 …… 63
一、组织培养细胞胞质 RNA 的制备 …… 64
二、胍盐法制备总 RNA …… 66
（一）氯化铯纯化培养细胞的 RNA …… 66
（二）氯化铯纯化组织中的 RNA …… 67
（三）一步法从培养细胞或组织中分离 RNA …… 68
三、细菌 RNA 的制备 …… 69
（一）从革兰阴性菌中分离高质量 RNA …… 69
（二）革兰阳性菌 RNA 的分离 …… 71
（三）革兰阴性菌 RNA 的快速分离 …… 71
四、poly(A)$^+$RNA 的制备 …… 72
第四节 核酸的定量测定 …… 73
一、紫外分光光度法测定 DNA 和 RNA 的含量 …… 73
二、电泳比较法（溴化乙锭荧光法） …… 74
三、定磷法测定核酸 …… 74
四、二苯胺法测定 DNA 含量 …… 75
五、地衣酚法测定 RNA 含量 …… 75
第三章 PCR 技术 …… 77
第一节 PCR 技术基础 …… 77
一、基本原理 …… 77
（一）反应过程 …… 77
（二）产物类型 …… 78
（三）平台效应 …… 78
二、PCR 技术的特点 …… 79
三、标准 PCR 反应 …… 79
四、影响因素 …… 80
（一）模板 …… 80
（二）引物 …… 81
（三）循环参数 …… 82
（四）*Taq* DNA 聚合酶及其浓度 …… 83
（五）Mg^{2+} 浓度 …… 83
（六）dNTP 浓度 …… 83
（七）其他辅助试剂 …… 84
五、扩增产物的检测分析 …… 84
（一）琼脂糖凝胶电泳 …… 84
（二）聚丙烯酰胺凝胶电泳 …… 85
（三）核酸探针杂交鉴定法 …… 85
（四）限制性内切酶分析法 …… 85
（五）单链构型多态性分析法 …… 86
六、PCR 条件优化 …… 86
（一）常规的优化方法 …… 86

（二）降落 PCR …… 86
七、PCR 常见问题与对策 …… 87
（一）PCR 污染 …… 87
（二）污染的预防 …… 87
（三）实验中的对照 …… 87
（四）扩增反应总是阴性结果（无产物）时应采取的措施 …… 87
（五）出现非特异性产物时应采取的措施 …… 88
第二节 PCR 技术的扩展 …… 88
一、逆转录 PCR …… 88
二、定量 PCR …… 89
（一）利用参照物的定量方法 …… 89
（二）竞争性定量 PCR …… 89
（三）荧光定量 PCR …… 90
三、重组 PCR …… 91
四、反向 PCR …… 92
五、不对称 PCR …… 92
六、复合 PCR …… 92
七、着色互补 PCR …… 93
八、锚定 PCR …… 93
九、原位 PCR …… 93
十、膜结合 PCR …… 93
十一、固着 PCR …… 93
十二、增效 PCR …… 94
第三节 PCR 技术的应用 …… 94
一、PCR 技术在分子生物学中的应用 …… 94
二、PCR 技术在传染病病原体检测中的应用 …… 95
三、PCR 在肿瘤相关基因检测中的应用 …… 96
四、PCR 技术在遗传病早期诊断中的应用 …… 97
五、PCR 在骨髓移植 HLA-D 位点配型中的应用 …… 98
六、PCR 技术在法医学鉴定中的应用 …… 98
七、PCR 在进化分析中的应用 …… 99
第四章 分子杂交与印迹技术 …… 100
第一节 核酸杂交的基本理论 …… 100
一、DNA 变性、复性及杂交 …… 100
二、核酸探针 …… 101
三、核酸杂交方法 …… 102
（一）影响杂交的因素 …… 102
（二）固相杂交法 …… 103
（三）液相杂交法 …… 104
（四）杂交信号的检测 …… 107
第二节 常用印迹杂交方法 …… 107

一、Southern 印迹杂交法 …… 107
二、Northern 印迹杂交法 …… 111
三、斑点及狭缝印迹杂交 …… 113
四、菌落原位杂交法 …… 113
五、核酸原位杂交法 …… 115
（一）组织与细胞的固定 …… 115
（二）载片和组织切片的处理 …… 115
（三）杂交 …… 116
（四）杂交后处理 …… 118
（五）显示 …… 119
（六）对照实验和核酸原位杂交结果的判断 …… 119
六、Western 印迹杂交法 …… 120
（一）蛋白质的分离 …… 120
（二）转膜 …… 120
（三）鉴定用膜的准备 …… 121
（四）鉴定 …… 122
第五章 分子克隆技术 …… 123
第一节 概述 …… 123
一、分子克隆的基本步骤 …… 123
二、分子克隆研究的内容 …… 124
（一）分子克隆工具的研究 …… 124
（二）分子克隆技术的研究 …… 124
（三）克隆对象——目的基因的研究 …… 124
（四）分子克隆产品的研究 …… 124
第二节 克隆操作中需要的工具酶 …… 125
一、限制性核酸内切酶 …… 125
（一）限制性核酸内切酶的类别 …… 126
（二）同工异源酶与同尾酶 …… 127
（三）限制性核酸内切酶活性及其影响因素 …… 128
二、DNA 连接酶 …… 128
三、DNA 聚合酶 …… 129
（一）DNA 聚合酶 Ⅰ …… 129
（二）Klenow 酶 …… 129
（三）*Taq* DNA 聚合酶 …… 129
（四）逆转录酶 …… 130
（五）T4 DNA 聚合酶 …… 130
（六）T7 DNA 聚合酶 …… 131
四、修饰酶 …… 131
（一）碱性磷酸酶 …… 131
（二）末端转移酶 …… 132
（三）T4 多核苷酸激酶 …… 132

（四）S1 核酸酶 …… 132
第三节　基因克隆载体 …… 133
一、质粒载体 …… 133
二、噬菌体载体 …… 135
三、黏粒载体 …… 137
第四节　基因克隆的一般方法 …… 137
一、获得目的基因的常用方法 …… 137
（一）化学合成法 …… 137
（二）逆转录 PCR 克隆目的基因 …… 138
（三）PCR 方法克隆目的基因 …… 139
（四）利用基因表达差异的克隆方法 …… 139
二、目的基因与载体连接 …… 141
（一）具黏性末端的 DNA 片段之间的连接 …… 141
（二）平末端的连接 …… 142
三、构建文库克隆目的基因 …… 143
（一）构建 cDNA 文库克隆目的基因 …… 143
（二）构建基因组文库克隆目的基因 …… 147
四、利用差异文库的克隆方法 …… 150
（一）差别杂交和扣除杂交克隆的目的基因 …… 150
（二）代表性差别分析技术克隆目的基因 …… 151
（三）抑制性扣除杂交技术 …… 153
第五节　克隆筛选与鉴定 …… 155
一、重组 DNA 导入受体细胞 …… 155
二、重组体的筛选与鉴定 …… 156
（一）根据遗传表型的筛选方法 …… 156
（二）依赖重组子结构特征的筛选法 …… 158
第六章　外源基因转移技术 …… 160
第一节　外源基因导入原核细胞的常用方法 …… 160
一、自然条件下的转化现象 …… 160
二、受体细胞的选择 …… 161
三、转化方法 …… 162
（一）$CaCl_2$ 诱导大肠杆菌转化法 …… 162
（二）电穿孔驱动的完整细胞转化 …… 164
（三）PEG 介导的细菌原生质体转化 …… 164
（四）通过接合作用传递质粒 DNA …… 165
（五）λ 噬菌体的转导和转染 …… 165
四、转化率及其影响因素 …… 166
第二节　外源基因导入真核细胞的方法 …… 167
一、物理方法 …… 168
二、化学方法 …… 170
三、生理方法 …… 174

第三节 病毒载体介导的基因转移 …… 175
一、RNA病毒载体及其介导的基因转移 …… 175
（一）逆转录病毒 …… 175
（二）逆转录病毒载体 …… 176
二、DNA病毒载体及其介导的基因转移 …… 179
（一）腺病毒载体 …… 179
（二）腺病毒相关病毒载体 …… 180
（三）疱疹病毒载体 …… 180
（四）嵌合（杂合）病毒载体 …… 181
第四节 报告基因的应用 …… 181
一、报告基因及其应用范围 …… 181
（一）报告基因 …… 181
（二）报告基因的应用 …… 182
二、常见的几种报告基因及其应用 …… 183
（一）荧光素酶 …… 183
（二）绿色荧光蛋白 …… 184
（三）氯霉素乙酰转移酶（CAT） …… 187
（四）β-半乳糖苷酶 …… 187
（五）分泌型的碱性磷酸酶（SEAP） …… 188
（六）人生长激素（hGH） …… 188
（七）β-葡糖醛酸酶（GUS） …… 189
第七章 蛋白质表达技术 …… 190
第一节 外源基因在原核细胞中的表达 …… 190
一、大肠杆菌表达系统 …… 190
（一）大肠杆菌表达载体的表达元件 …… 190
（二）大肠杆菌的表达载体 …… 194
（三）外源基因在大肠杆菌中表达 …… 197
（四）提高外源基因表达水平的措施 …… 199
（五）包含体 …… 200
二、芽孢杆菌表达系统 …… 201
（一）芽孢杆菌基因表达的特点 …… 201
（二）芽孢杆菌表达系统 …… 202
（三）存在的问题 …… 203
三、链霉菌表达系统 …… 204
（一）链霉菌的生物学特征 …… 204
（二）链霉菌基因表达的特点 …… 205
（三）链霉菌基因表达系统 …… 206
（四）影响链霉菌中基因表达的因素 …… 208
（五）链霉菌表达系统的优缺点 …… 209
第二节 外源基因在真核细胞中的表达 …… 209
一、酵母表达系统 …… 209

(一) 酵母表达系统的组成 …… 210
(二) 酵母表达系统的应用策略 …… 213
二、昆虫表达系统 …… 216
(一) 杆状病毒表达系统 …… 216
(二) 家蚕核型多角体病毒 (BmNPV) 表达系统 …… 220
第三节 哺乳动物细胞表达系统 …… 220
一、哺乳动物细胞表达系统的构成 …… 221
(一) 宿主细胞 …… 221
(二) 表达载体 …… 222
(三) 目的基因 …… 225
二、常用的细胞表达系统 …… 225
三、基因的导入方法 …… 226
四、外源基因在哺乳细胞的表达和基因表达产物的检测 …… 226
五、哺乳动物细胞表达系统的改进策略 …… 226
(一) 改进表达载体，提高表达水平 …… 226
(二) 改造宿主细胞及培养手段 …… 228
(三) 提高产品质量 …… 229
(四) 筛选高表达细胞克隆的新方法 …… 230
第四节 重组蛋白的检测与鉴定 …… 230
一、定量分析 …… 230
二、纯度分析 …… 230
三、重组蛋白质的鉴定 …… 231
第八章 分子标记技术 …… 233
第一节 概述 …… 233
一、放射性同位素 …… 233
(一) 放射性同位素的基本性质 …… 233
(二) 放射性强度及其度量单位 …… 234
(三) 射线与物质的相互作用 …… 234
(四) 射线的防护 …… 234
二、放射性同位素标记 …… 235
三、非同位素标记 …… 236
(一) 发光信号的特点 …… 236
(二) 化学发光的本质 …… 237
(三) 生物发光 …… 237
(四) 荧光的基础 …… 237
第二节 核酸的标记 …… 238
一、放射性同位素标记探针 …… 238
(一) 切口平移法 …… 238
(二) 末端标记法 …… 240
(三) 随机引物法标记 DNA 片段 …… 243
(四) 聚合酶链反应标记高活性 DNA 探针 …… 244

二、非放射性同位素标记 …… 244
（一）间接非同位素标记 …… 245
（二）用生物素、地高辛或荧光素标记探针的方法 …… 245
（三）DIG 标记实例 …… 248
（四）直接非同位素标记 …… 249
第三节 蛋白质的标记 …… 250
一、荧光标记 …… 250
（一）利用氨基的标记方法 …… 250
（二）利用巯基的标记方法 …… 252
（三）利用醇羟基的标记方法 …… 253
二、利用酶、酶的底物及酶抑制剂的标记 …… 253
（一）糖苷酶及其底物 …… 254
（二）磷酸酶类及其底物 …… 254
（三）其他酶与底物 …… 255
（四）生物素与亲和素系统 …… 255
三、利用抗体的标记技术 …… 256
（一）免疫荧光标记技术 …… 256
（二）放射免疫标记技术 …… 257
（三）酶免疫标记技术 …… 258
（四）胶体金标记技术 …… 258
（五）发光免疫标记技术 …… 259
第九章 分子改造技术 …… 261
第一节 概述 …… 261
第二节 蛋白质的体外改造 …… 262
一、寡核苷酸指导的诱变方法 …… 262
（一）非 PCR 的诱变方法 …… 263
（二）PCR 方法 …… 263
（三）诱变寡核苷酸的设计 …… 265
二、随机诱变 …… 266
（一）化学诱变 …… 266
（二）丙氨酸扫描诱变 …… 267
三、突变体筛选方法 …… 267
（一）限制酶切割区分亲本模板与突变产物 …… 268
（二）单一限制位点的消除 …… 268
四、分子进化 …… 268
（一）概念 …… 268
（二）基本方法 …… 269
第三节 蛋白质改造技术的应用 …… 271
一、点突变 …… 271
二、结构域改组 …… 271
三、全蛋白的重组 …… 272

四、蛋白质-配体相互作用 …… 272
第四节 单克隆抗体的改造 …… 273
一、抗体的产生和抗体的结构 …… 274
二、多抗、单抗及重组抗体 …… 275
三、抗体改造 …… 276
（一）人-鼠嵌合抗体 …… 276
（二）小分子抗体 …… 277
（三）双功能抗体 …… 277
（四）抗体融合蛋白 …… 278
第十章 测序及人工合成技术 …… 279
第一节 DNA 序列测定 …… 279
一、Sanger 双脱氧链终止法 …… 279
（一）Sanger 双脱氧链终止法的原理 …… 279
（二）Sanger 双脱氧链终止法所需关键试剂 …… 280
（三）用于测序的变性凝胶电泳 …… 282
（四）大片段 DNA 的测序策略 …… 282
二、Maxam-Gilbert 化学修饰法 …… 283
（一）化学裂解法测序的原理 …… 283
（二）化学试剂特异断裂 DNA 的机制 …… 284
三、DNA 序列的自动测序 …… 284
第二节 DNA 的化学合成 …… 285
一、DNA 化学合成的原理 …… 286
（一）磷酸三酯法合成寡核苷酸的原理 …… 286
（二）固相亚磷酸三酯法合成寡核苷酸的原理 …… 286
（三）用寡核苷酸片段组装基因的方式 …… 288
二、DNA 合成技术在分子生物学研究中的应用 …… 289
第三节 蛋白质和多肽的氨基酸序列测定 …… 291
一、序列测定前的准备 …… 291
（一）蛋白质纯化 …… 291
（二）肽链的分离 …… 292
（三）肽链的部分裂解 …… 292
（四）肽片段的分离 …… 293
二、序列测定的方法和原理 …… 294
（一）手工测序 …… 294
（二）自动序列分析 …… 297
三、影响测序的因素 …… 298
第四节 蛋白质或多肽的化学合成 …… 298
一、氨基酸官能团的保护 …… 299
二、羧基的活化和肽键的生成 …… 301
三、合成肽的脱保护基及纯化 …… 302
第十一章 基因组学技术 …… 305

第一节　结构基因组学及常用研究技术 …… 305
一、基因组作图 …… 305
二、新基因的分离——cDNA 末端快速扩增（RACE）技术 …… 307
三、基因组序列测定技术 …… 310
四、基因定位技术 …… 311
（一）荧光原位杂交（FISH）技术 …… 311
（二）辐射杂种细胞系（RH）技术 …… 312
第二节　功能基因组学及常用研究技术 …… 312
一、基因突变检测（分析）技术 …… 313
（一）单链构象多态性（SSCP）技术 …… 313
（二）变性梯度凝胶电泳（DGGE）技术 …… 315
（三）直接测序法 …… 317
（四）单碱基延伸标签阵列技术（SBE-TAGS） …… 318
（五）异源双链分析技术 …… 318
（六）连接酶链反应 …… 318
（七）等位基因特异性寡核苷酸杂交（ASOH） …… 318
（八）RNA 酶 A 切割法 …… 319
（九）基于 PCR 与酶切的技术 …… 319
（十）高通量检测技术 …… 320
二、比较基因组杂交技术 …… 321
三、微阵列-比较基因组杂交技术 …… 323
四、染色体原位杂交 …… 325
五、基因表达分析技术 …… 325
（一）基因表达系列分析技术（SAGE） …… 326
（二）RNase 保护试验（RPA） …… 328
六、模式生物体研究 …… 328
（一）转基因动物 …… 329
（二）基因打靶技术 …… 334
（三）时空可调节性基因打靶技术 …… 337
（四）基因陷阱 …… 338
（五）诱变技术在功能基因组学中的应用 …… 341
七、SNP、EST 在研究新（未知）基因中的应用 …… 341
（一）单核苷酸多态性（SNP） …… 341
（二）表达序列标签（EST） …… 343
第十二章　蛋白质组学技术 …… 345
第一节　概述 …… 345
一、蛋白质组学研究的意义和背景 …… 345
二、蛋白质组学研究的特点 …… 346
第二节　蛋白质组学研究的技术体系与路线 …… 350
一、蛋白质组分离技术 …… 350
（一）双向凝胶电泳的特点 …… 350

（二）双向凝胶电泳的基本步骤 …… 351
（三）蛋白质组分离的非凝胶技术 …… 356
二、蛋白质组分析技术 …… 358
三、蛋白质组数据库的建立和生物信息学 …… 359
四、蛋白质组研究的发展趋势及应用前景 …… 359
第三节 蛋白质组学技术平台与完整解决方案 …… 360
一、蛋白质组学技术平台 …… 360
二、蛋白质组学研究的完整解决方案 …… 361
第十三章 生物芯片技术 …… 364
第一节 概述 …… 364
一、生物芯片的概念和特点 …… 364
二、生物芯片的分类 …… 365
第二节 基因芯片 …… 366
一、基本原理 …… 366
二、基因芯片的制备和检测 …… 366
（一）芯片载体的表面修饰 …… 366
（二）芯片阵列制作方式 …… 367
（三）杂交探针的种类与制备 …… 368
（四）样品制备和标记 …… 370
（五）杂交 …… 370
（六）杂交信号采集和处理 …… 371
（七）生物信息的分析和提取 …… 371
第三节 蛋白质芯片 …… 373
一、基本原理 …… 374
二、蛋白质芯片制备和检测的基本流程 …… 377
（一）蛋白质芯片制备与检测流程 …… 377
（二）待检样品的制备和标记 …… 379
（三）蛋白质芯片的杂交反应和结果检测 …… 380
三、蛋白质芯片的研究进展 …… 381
（一）方法学研究和新技术的应用 …… 381
（二）抗体芯片 …… 382
第四节 生物芯片在医学中的应用 …… 382
一、基因芯片在医学中的应用 …… 382
（一）微缩实验室芯片 …… 382
（二）基因突变分析 …… 383
（三）克隆选择及文库筛选 …… 383
（四）测序芯片 …… 384
（五）基因多态性分析 …… 384
（六）基因表达谱分析 …… 384
（七）微生物菌种鉴定、致病机理及耐药性分析 …… 385
（八）药物筛选芯片 …… 385

二、蛋白质芯片在医学中的应用 …… 385
(一) 蛋白质组-基因组链接 …… 385
(二) 高通量的蛋白质表达筛选及药物研究 …… 386
(三) 蛋白质和蛋白质之间相互作用的研究 …… 386
(四) 在疾病诊断方面的应用 …… 388
(五) 展望 …… 388
第十四章 生物信息学技术 …… 390
第一节 概述 …… 390
一、生物信息学的诞生和发展 …… 390
二、生物信息学研究的基本方法 …… 390
三、生物信息学的目的意义 …… 391
第二节 常用核酸序列数据库 …… 392
第三节 常用蛋白质序列数据库 …… 398
第四节 序列分析 …… 402
一、序列比对方法和相似性搜索 …… 403
二、FASTA …… 404
三、BLAST …… 405
第五节 蛋白质结构预测 …… 410
一、二级结构和折叠类型 …… 410
二、三级结构预测 …… 411
第十五章 RNA 研究技术 …… 414
第一节 RNA 结构分析和合成分析 …… 414
一、用单链 DNA 探针和 S1 核酸酶分析 mRNA …… 414
二、核糖核酸酶保护测定 …… 419
三、引物延伸法 …… 422
四、RNA run-on 分析 …… 424
第二节 RNA 组学技术 …… 425
一、反义核酸技术 …… 426
(一) 反义 RNA 作用机制 …… 426
(二) 反义 RNA 技术方法 …… 427
(三) 反义 RNA 技术应用 …… 428
二、RNA 干扰 (RNAi) 技术 …… 428
(一) RNAi 作用的机制 …… 429
(二) RNAi 实验方案 …… 430
(三) RNAi 敏感性的预测 …… 432
(四) siRNA 导入宿主细胞 …… 433
(五) 结果分析 …… 434
(六) RNAi 在基因功能分析中的应用 …… 434
第三节 核酶技术 …… 434
第四节 RNA 错折叠 (ODMiR) 技术 …… 436
附录 …… 438
参考文献 …… 451

第一章 生物大分子制备和分析常用技术

生物大分子主要包括蛋白质、核酸、多糖及部分脂类，而分子生物学中研究的主要对象是蛋白质与核酸。以蛋白质和核酸的结构与功能为基础，从分子水平上认识生命现象，已经成为现代生物学发展的主要方向。研究生物大分子，首先要得到高度纯化并具有生物活性的目标物质。生物大分子的制备一般分为以下几个阶段：预处理，细胞的分离，细胞的破碎，细胞器的分离，生物大分子的提取、纯化、浓缩、干燥及保存。本章将主要讨论以蛋白质和核酸为代表的生物大分子的制备和分析常用技术。

第一节 概 论

生物大分子的分离纯化方法很多，其原理主要是利用各分子之间特异性的差异，如分子大小、形状、酸碱性、溶解度、极性、电荷及亲和性等。生物大分子的制备工作涉及物理、化学和生物学等方面的知识。各种分离纯化方法的主要原理可归纳为两个方面：一是利用混合物中各组分之间分配系数的差异，将它们分配到两个物相或几个物相中，如盐析、盐溶、分配色谱、有机溶剂沉淀、共沉淀和结晶等；二是将混合物置于单一的物相中，通过物理力场的作用使各组分分配于不同区域而达到分离的目的，如电泳、离心、超滤等。目前纯化蛋白质的 3 种关键技术是电泳、色谱和离心。表 1-1 将生物大分子制备的方法按照分子的大小、形状、密度、溶解度、带电性质及亲和力等主要因素进行了分类。在实际工作中，往往需要综合使用多种方法才能制备出一种纯化的生物大分子。在所有这些方法的应用中必须注意保存生物大分子的完整性，防止酸、碱、高温、剧烈机械作用而导致分离产物生物活性的丧失。

表 1-1 生物大分子分离纯化方法类型

性 质	具 体 方 法
分子大小和形态	离心、超滤、分子筛、透析、凝胶色谱、SDS-聚丙烯酰胺凝胶电泳
分子的密度	超速离心
溶解度	盐析、有机溶剂沉淀、等电点沉淀、分配色谱等
电荷差异	电泳(醋酸纤维素薄电泳膜、琼脂糖凝胶电泳、聚丙烯酰胺凝胶电泳)、等电聚焦电泳、离子交换色谱
生物功能专一性	亲和色谱

一、预处理和细胞的分离

（一）选择材料及预处理

微生物、植物和动物都可作为制备生物大分子的原材料，所选用的材料主要依据实验目的来确定。

对于微生物原料，应注意它的生长期，例如在微生物的对数生长期，酶和核酸的含量较高，可以获得高产量。以微生物为材料时，生物大分子可分为两种情况：①微生物菌体分泌到培养基中的代谢产物和胞外酶等；②菌体含有的生化物质，如蛋白质、核酸和胞内酶等。

植物材料必须经过去壳、脱脂，应注意植物品种和生长发育状况不同，其中所含生物大分子的含量变化很大，并与季节关系密切。

对动物组织，要选择有效成分含量丰富的脏器组织为原材料。同时要注意动物的年龄、性别、种类等影响因素。

各类生物大分子在细胞内的分布是不同的，如 DNA 几乎全部在细胞核内，RNA 则主要在胞浆，各种酶在细胞内的分布也有特定的位置。因此，应根据某一目的物质的位置来选取材料。

不同来源的样品，其预处理过程应有所不同，但都应做到及时、科学地处理。对动物材料，应先去除结缔组织、脂肪组织等非活性部分，然后选择合适的细胞破碎方法处理样品，并用相应的溶剂提取。植物组织要先去壳、去皮、除脂，只保留所需要的部分。微生物材料要及时将菌体与培养液分离。如果对预处理好的材料不立即进行实验，应将其冷冻保存；而分离易分解的生物大分子应选用新鲜材料。

（二）细胞的分离

分离细胞是生物学实验中的基本技术之一，它主要根据细胞本身的某些性质来分离具有同一性状的细胞群，这些性质包括：细胞大小、细胞密度、细胞表面电荷、细胞表面标志、每个细胞的散射光线总量、细胞中一个或多个成分的荧光、细胞对其他介质的吸附作用。选用细胞分离技术时，除根据不同目的而选用外，原则上要求所用方法既简便可行，又能获得高纯度、高收获量、高活力的细胞。以下是几种常用的细胞分离方法。

1. 梯度沉降分离法

梯度沉降分离法主要根据细胞大小分离细胞。细胞在单位重力作用下，通过密度介质，或在低离心力作用下，通过梯度密度溶液沉降。由于细胞大小不同，沉降速度不同，细胞越大，沉降越快。

在分离细胞过程中，为了稳定沉降细胞，需要使用适当的分离介质形成一定的梯度密度溶液，常用的分离介质有血清、聚蔗糖和蔗糖等。本法常用于分离大小差异较明显的培养细胞。

2. 等密度沉降分离法

等密度沉降分离法主要根据细胞密度差异分离细胞。细胞在连续密度梯度分离介质中，在强离心力作用下，细胞最后到达与其密度相同的分离介质层面，并能保持平衡。在非连续密度梯度中，分离细胞主要集中于介于其自身密度的两种密度介质交界面上，从而达到分离目的。

目前常用于细胞分离的介质有白蛋白、Ficoll（一种蔗糖多聚体）、Percoll、Metrizamide。这里仅以 Percoll 为例作一说明。Percoll 是商品名，它是一种外面被聚乙烯吡咯烷酮包被的硅胶颗粒。这种密度梯度分离剂无毒，无刺激性，对细胞无吸附作用，产生的渗透压很小，因此在全部密度范围保持相等的张力。Percoll 在离心过程中会自然形成密度梯度。Percoll 的基本性质：密度为 (1.130 ± 0.005)g/ml，黏度为 $(10\pm5)\times10^{-3}$Pa·s，渗透压 <20mOsm/kg H_2O，pH 为 8.9 ± 0.3。近年用 Percoll 对多种细胞进行了分离，如人血细胞、人骨髓细胞、鼠肝细胞等。

3. 流式细胞仪分离法

本法是以免疫荧光法使荧光抗体与细胞膜表面抗原结合，继而用超声波使其充分分散为单个细胞，再将细胞悬液以 10m/s 速度通过一个直径 50μm 的喷嘴喷出，使细胞悬液成为极细的微滴，每微滴至多含一个细胞。经激光照射时，细胞上所带的荧光被激发而转换成脉冲，测定脉冲数可换算出各种不同细胞表面抗原的情况。同时，由于悬浮微滴中带不同程度负电荷的细胞，受电场影响后移行偏斜程度不同，而不带电荷的细胞仍以直线流过，为此可收集到不带电荷或带电荷多少不同的各种细胞，即为不同群的细胞。本法优点是分离速度很快、分离的细胞仍保持各种功能。

二、细胞的破碎及细胞器的分离

目前已建立很多破碎细胞、释放细胞内容物的方法。根据作用方式不同，基本可以分为两大类：机械法和非机械法。传统的机械法包括匀浆、研磨、压榨、超声等；常见的非机械法包括渗透、酶溶、冻融、化学等。另有一些新的方法也在不断发展和完善，如激光破碎、冷冻喷射、相向流撞击等。每一种方法都有它的适用范围及优缺点，选择的标准常常要依靠过去经验和预实验情况。选择的原则是该方法不影响目的蛋白或产物的结构和功能，或者说，采用的方法可以避免造成目的蛋白或产物的变性或失活。常见组织和细胞的破碎和溶解一般都采用传统的标准方法；有时为了提高产物的回收率或针对产物的特性，尝试新的细胞筛选、破碎及溶解的方法也很有必要。

细胞破碎前，一般用缓冲液洗去组织样品中残留的血液和污染物；培养细胞通常用缓冲液混悬后离心，除去残留培养液等。细胞破碎获得的抽提物称为匀浆。匀浆根据不同需要可以在 12000*g* 离心 10min 至 100000*g* 离心 1h 范围内离心。沉淀主要含有膜组分等，弃去。上清液含有目的蛋白，称为粗提物。如果粗提物有漂浮颗粒，可用纱布或玻璃纤维滤去后再进一步纯化。各种组织和细胞的常用裂解方法列于表 1-2。

表 1-2　各种组织和细胞的破碎方法

细胞破碎方法	组织种类	细胞破碎方法	组织种类
旋刀式匀浆	大多数动、植物组织	酶溶	细菌、酵母
手动式匀浆	柔软的动物组织	去垢剂渗透	组织培养细胞
超声破碎	细胞混悬液	有机溶剂渗透	细菌、酵母
高压匀浆	细菌、酵母、植物细胞	低渗裂解	红细胞、细菌
研磨	细菌、植物细胞	冻融裂解	培养细胞
高速珠磨	细胞混悬液		

1. 匀浆法

匀浆是机体软组织破碎最常用的方法之一。它的工作原理是通过固体剪切力破碎组织和细胞，释放蛋白进入溶液。市售的匀浆器主要有 4 类：刀片式组织破碎匀浆器、内切式组织匀浆器、玻璃匀浆器和用于规模生产的高压匀浆器。玻璃匀浆器的匀浆杵有玻璃制的，也有特夫隆（Teflon）制的，既可以手动也可以电动。由于匀浆过程中蛋白质被蛋白酶降解的可能性较小，所以匀浆是简便、迅速和风险小的组织破碎方法，是实验室首先考虑的方法之一。

匀浆过程应注意维持低温。玻璃匀浆器可外置冰水浴，其他匀浆容器可预冷或在冷室内匀浆。匀浆时所需匀浆缓冲液的体积在不同条件下可有很大差别，有时甚至可为湿重组织体

积的9～10倍。

（1）杵状玻璃匀浆器法　该匀浆器由一根一端表面磨砂的玻璃杵和一个内壁磨砂的玻璃管组成。使用时，先用锋利的刀片把组织切成碎块，然后把组织碎块或细胞悬液加入套管后，手工或以电动搅拌机旋转研磨，玻璃杵在套管中上下移动时产生机械切力使细胞破碎。这种方法对生物大分子破坏少，是目前广泛使用的一种细胞破碎方法。

（2）高速组织匀浆器法　匀浆器由调速器、支架、马达、带杆叶片刀和梅花玻璃杯组成。使用时先将4℃预冷的组织碎块或细胞悬液加入玻璃杯中，至杯体积的1/3即可，盖好玻璃杯盖，固定好带杆叶片刀，缓缓调节旋转速度。一般开机数十秒后，组织细胞即可被高速旋转的叶片刀破碎。组织匀浆器高速转动时易产热，可导致分离物的降解，因此注意不能持续时间过久，必要时使用循环冷却水降温。

2. 超声波法

输入高能超声波可以破碎细胞，其机理可能是与超声波作用溶液时气泡产生、长大和破碎的空化现象有关。空化现象引起的冲击波和剪切力使细胞裂解。超声破碎的效率取决于超声波频率、功率、处理时间、细胞浓度及细胞类型等。

使用超声破碎必须注意的是控制强度在一定限度，即刚好低于溶液产生泡沫的水平。因为产生泡沫会导致蛋白质变性。过低的强度将降低破碎效率。最好在正式实验前用多余样品试验，调校超声波发生器在稍低于产生泡沫的强度。正式超声样品时强度只能在预定位置附近做微小调整。

超声破碎在处理少量样品时操作简便，液量损失少；此法多适用于微生物材料和脑组织匀浆。用大肠杆菌制备各种酶，常选用菌体质量浓度为50～100mg/ml，在10～100kHz频率下处理10～15min。但是，此法的缺点是超声波产生的化学自由基团能使敏感的活性物质变性失活。大容量装置声能传递、散热均有困难，在处理过程中会产生大量的热，应采取相应降温措施。对超声波敏感的酶和核酸应慎用。

3. 高压匀浆法

此方法的原理是细胞悬液从高压室的环状隙高速喷射到静止的撞击环上，被迫改变方向经出口管流出，此过程中细胞经历了高速造成的剪切、碰撞及高压到常压的变化，从而破碎释放内含物。增加压力或增加破碎次数都能提高破碎效率，但压力增加到一定程度零部件磨损较大。通常撞击环较易磨损，应定期更换。为防止污染，使用前后高压室至出口管需彻底清洗。

该法常用来裂解细菌和酵母。适用于中等体积（10～30ml）的样品，体积太大或太小均不适宜。

4. 研磨法

研磨是破碎单一细胞的有效措施。借助研磨中磨料和细胞间的剪切及碰撞作用破碎细胞。常用的磨料为沙子、氧化铝。在研钵内样品与磨料被研磨成稠糊状。一次破碎的细胞量湿重可达30g。

此法主要用于细菌、酵母和一些植物组织的破碎。研磨时间一般不超过15min。石英沙或氧化铝用前需做清洁处理。研磨时产生劈啪声提示可能效果不错。

5. 高速珠磨法

该法应视为研磨法的扩展。它用玻璃珠替代磨料。小量样品（湿重不超过3g）可在试管内进行，大量样品需使用特制的高速珠磨机。

本法最常用来破碎酵母。使用的玻璃珠直径 500μm，用前需如下处理：浓盐酸洗后，双蒸水洗至中性，烘干，用前预冷。宜用有密封圈的螺旋盖关紧试管或用石蜡膜封好，以防混旋时液体漏出。

6. 酶溶法

酶溶法就是用生物酶将细胞壁和细胞膜消化溶解的方法。因此利用此方法处理细胞必须根据细胞的结构和化学组成选择适当的酶。常用的有溶菌酶、β-1,3-葡聚糖酶、β-1,6-葡聚糖酶、蛋白酶、甘露糖酶、糖苷酶、肽链内切酶、壳多糖酶等。细菌主要用溶菌酶处理，酵母需用几种酶进行复合处理。使用溶酶系统时要注意控制温度、酸碱度、酶用量、酶的先后次序及处理时间。

虽然酶溶法有选择性释放产物、核酸泄出量少、细胞外形完整等优点，但它也存在明显不足，如酶价格高、通用性差、有时存在产物抑制等。

7. 化学渗透法

有些有机溶剂（如苯、甲苯）、抗生素、表面活性剂（SDS、NP-40、曲通 X-100 等）、金属螯合剂（EDTA）、变性剂（盐酸胍、脲）等化学药品都可以改变细胞壁或细胞膜的通透性，使内含物有选择地渗透出来。

此法主要用于溶解培养的动物细胞。低浓度表面活性剂处理，如能充分溶解细胞，则是很温和的方法。若目的蛋白不稳定，冰浴中孵育时间可以减少甚至取消；裂解液中加入 0.2 倍体积 50%甘油也可以稳定目的蛋白。

8. 反复冻融法

将细胞在－20℃以下冰冻，室温融解，反复几次，由于细胞内冰粒形成和剩余细胞液的盐浓度增高引起溶胀，使细胞结构破碎。

它仅适用于提取非常稳定的蛋白。对于韧性很强的组织，如皮肤、肌腱等，可用液氮冻硬变脆，敲碎成小块后，在研钵中加液氮研成粉，再加缓冲液溶解。

9. 低渗裂解法

此法是无细胞壁细胞在低渗溶液中通过渗透张力作用裂解的方法，常用于红细胞的裂解。

无论用哪一种方法破碎组织细胞，都会使细胞内蛋白质或核酸水解酶释放到溶液中，使大分子物质降解，导致天然物质量的减少。加入二异丙基氟磷酸（DFP）可以抑制或减慢自溶作用；加入碘乙酸可以抑制那些活性中心需要有巯基的蛋白水解酶的活性；加入苯甲磺酰氟化物（PMSF）能消除蛋白水解酶的部分活力，但不是全部；还可通过选择 pH、温度或离子强度等，使这些条件适合于目的物质的提取。

生物大分子的分析技术众多，其中以电泳技术、色谱技术、离心技术及各种光谱技术等应用最为普遍。本章将主要就这些常用技术做详细介绍。

第二节　电泳技术

电泳是指带电粒子在电场的作用下，在一定介质（溶剂）中所发生的向异性电极定向移动的现象。利用电泳现象对混合物进行分离、分析的技术叫做电泳技术。

电泳的分类方法很多，根据有无支持物，可分为区带电泳和自由电泳。区带电泳根据所用支持物不同常有不同的名称，例如纸电泳、醋酸纤维素薄膜电泳、琼脂糖凝胶电泳、聚丙烯酰胺凝胶电泳等。另外，根据电压的高低分为低电压电泳和高电压电泳；根据支持物的形状分为薄层电泳、平板电泳（水平平板电泳，垂直平板电泳）、圆盘电泳、毛细管电泳；根据用途分为分析电泳、制备电泳、定量免疫电泳；根据区带形式分为盘状电泳、火箭电泳等。

电泳技术具有以下特点：①几乎适用于所有带电物质的分离，并可进行定性或定量分析；②样品用量极少；③设备简单；④可在常温进行；⑤操作简便省时；⑥分辨率高。目前电泳技术已经广泛应用于基础理论研究、临床诊断及工业制造等方面，用于分离酶、蛋白质、核酸等大分子的研究工作，对生物化学与分子生物学的发展起了重要作用。特别是双向凝胶电泳和质谱技术的建立和联用，提供了蛋白质组研究所需的技术体系，使得蛋白质组的研究成为可能。

一、基本原理

任何物质由于其本身的解离作用或表面上吸附其他带电质点，在电场中便会向一定的电极移动。作为带电颗粒可以是小的离子，也可是生物大分子，如蛋白质、核酸、病毒颗粒、细胞器等。

设一带电粒子在电场中所受的力为 F，F 的大小决定于粒子所带电荷 Q 和电场强度 E，即：

$$F=QE$$

根据斯托克斯定律（Stoke's law），一球形的粒子运动时所受到的阻力 F'，与粒子运动的速度 v，粒子的半径 r，介质的黏度 η 的关系为：

$$F'=6\pi r\eta v$$

当电泳达到平衡，带电粒子在电场做匀速运动时，则：

$$F=F'$$

即：

$$QE=6\pi r\eta v$$

移项得：

$$\frac{v}{E}=\frac{Q}{6\pi r\eta}$$

v/E 表示粒子在单位电场强度时的平均电泳速度，称为电泳迁移率（electrophoretic mobility），用 μ 来表示。

$$\mu=\frac{v}{E}=\frac{d}{t\cdot E}=\frac{Q}{6\pi r\eta}$$

可见一个带电粒子的迁移率不仅取决于粒子本身的所带电荷的多少，还取决于粒子的大小、形状、介质的黏度等多种因素。

由于不同物质颗粒的大小和所带电荷往往各不相同，所以两种不同的颗粒（如两种蛋白质分子）通常有不同的迁移率。假设某物质 A 和物质 B 在同一电场中移动的距离为：

$$d_A=\mu_A\cdot Et \quad d_B=\mu_B\cdot Et$$

则两物质移动距离之差为：

$$\Delta d=(d_A-d_B)=(\mu_A-\mu_B)\cdot Et$$

上式指出物质A、B能否利用电泳进行分离，决定于两者迁移率的差别，有差别时才能分离，如两者的迁移率相同则不能被分离。两物质电泳后分开的距离与两者的迁移率差、电泳时的电压、电泳时间等因素相关。

二、影响电泳的因素

影响电泳结果的因素很多，下面讨论5个主要方面。

1. 样品的性质

被分离颗粒带电荷的多少、分子的大小、形状影响电泳的速度。电泳速度与颗粒带电荷的多少成正比，与分子量大小成反比，球形的分子要比纤维状的分子泳动速度快。

2. 电场强度

电场强度是每厘米的电位降，亦即电势梯度。电场强度愈高，带电质点移动速度也愈快。

3. 缓冲液

缓冲液能使电泳中的支持介质保持稳定的pH，并通过它的组成成分、浓度等因素影响着颗粒的迁移率。

（1）pH　溶液的pH决定带电颗粒解离程度，是颗粒带有净电荷多少的决定因素。对蛋白质、氨基酸等两性电解质来说，缓冲液的pH距等电点（p*I*）越远，所带净电荷就越多，电泳速度也越快；反之则越慢。因此，当分离蛋白质溶液时，应选择一个合适的pH，使各种蛋白质所带净电荷的差异量增大，以利于分离。通常进行血清蛋白电泳时，当采用pH8.6的缓冲液时，pH大于血清中各种蛋白质的等电点，蛋白质均带负电荷，电泳时向正极移动。

（2）成分　通常采用的是甲酸盐、乙酸盐、柠檬盐、磷酸盐、巴比妥盐和三羟甲基氨基甲烷-乙二胺基四乙酸缓冲液等。要求缓冲液的物质性能稳定，不易电解。血清蛋白分离时最常用的是巴比妥-巴比妥钠组成的缓冲液。

（3）浓度　缓冲液的浓度可用摩尔浓度或离子强度表示。离子强度低，电泳速度快，分离区带不易清晰，如果过低，缓冲液的缓冲量小，难以维持pH的恒定；离子强度高，电泳速度慢，但区带分离清晰，如果过高，除了使得电泳速度过慢，还会因为离子强度增加使电泳时的电流变大，增加热量的产生，引起温度的改变，这对电泳是很不利的。所以对缓冲液离子强度的选择，必须两者兼顾，一般是在0.02～0.2mol/L之间。

离子强度的计算公式为 $I=1/2\sum c_i Z_i^2$，式中 I 为离子强度；c_i 为离子的摩尔浓度；Z_i 为离子的价数；$\sum$ 表示“加和”，例如0.15mol/L NaCl的离子强度为：

$$I=1/2(0.15\times 1^2+0.15\times 1^2)=0.15$$

4. 支持介质

支持介质要求是应具较大惰性的材料，且不与被分离的样品或缓冲液起化学反应。此外，还要求具有一定的坚韧度，不易断裂，容易保存。由于各种介质的精确结构对一种被分离物的移动速度有很大影响，所以对支持介质的选择应取决于被分离物质的类型。

（1）吸附　支持介质的表面对被分离物质具有吸附作用，使分离物质滞留而降低电泳速度，会出现样品的拖尾。由于介质对各种物质吸附力不同，因而降低了分离的分辨率。纸的吸附性最大，醋酸纤维素薄膜的吸附作用很小。

（2）电渗　当固体与溶液接触时，许多固体表面都带有多余（额外）的负电荷，溶液中的正离子吸附至固体表面形成双电子层，当在液体两端施加了电压时，就会发生液体相对于

固体表面的移动，这一现象称为电渗（图 1-1），电渗中液体的整体流动叫电渗流。如纸电泳时，滤纸中含有的羟基使表面带负电荷，与表面接触的水溶液则带正电荷，溶液向负极移动。

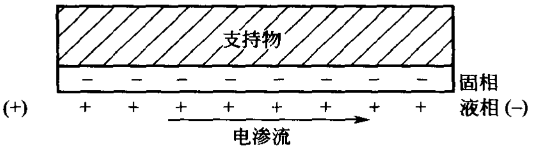


图 1-1　电渗作用示意图

由于电渗现象与电泳同时存在，所以电泳时分离物质的电泳速度也受电渗的影响。例如血清蛋白电泳，在巴比妥盐缓冲液 pH8.6、离子强度 0.06 的条件下进行，蛋白质的移动方向与电渗现象的水溶液移动方向相反，蛋白质泳动的距离等于电泳泳动距离减去电渗的距离，使电泳速度减慢。滤纸、淀粉胶、琼脂糖凝胶的电渗现象较大，醋酸纤维素、聚丙烯酰胺凝胶等介质的电渗现象要小得多。

（3）分子筛　有的支持介质具有分子筛作用，如凝胶。凝胶的分子筛效应使得样品增加了按分子量大小进行分离的作用，提高了电泳的分辨率。

5. 温度

电泳时电流通过支持介质会产生热量，按焦耳定律，电流通过导体时的产热与电流强度的平方、导体的电阻和通电的时间成正比（$Q=I^2Rt$）。产生热量对电泳技术是不利的，因为产热可促使支持介质上溶剂的蒸发，而影响缓冲溶液的离子强度。温度升高时，介质黏度下降，分子运动加剧，引起自由扩散变快，迁移增加。温度每升高 1℃，迁移率约增加 2.4%。若温度过高可导致分离样品变性而使电泳失败。为降低热效应对电泳的影响，可控制电压或电流，也可在电泳系统中安装冷却散热装置。对高压电泳必须增设冷却系统，以防样品在电泳时变性。

三、醋酸纤维素薄膜电泳

醋酸纤维素薄膜（cellulose acetate membrane，CAM）是一种由醋酸纤维素加工制成的细密且薄的微孔膜。根据其乙酰化程度、厚度、孔径和网状结构等方面不同而具有不同类型，现已广泛应用于医学临床中各种生物分子的分离、分析中，如分离血清蛋白、血红蛋白、球蛋白、脂蛋白、糖蛋白、甲胎蛋白、类固醇及同工酶等。它具有简单、快速等优点，电渗作用均一，不影响样品的分离效果；不足之处是分辨率比聚丙烯酰胺凝胶电泳低，由于薄膜厚度小（约 10～100μm），样品用量很少，不适于制备。

（1）影响电泳的因素　①膜的性质；②缓冲溶液的组成、pH 和浓度；③电流和有效电压；④电泳时间和温度；⑤电泳槽的密封性能。

（2）缓冲液的选择　为了提高电泳速度和分辨率，应根据样品的理化性质选择缓冲液的种类、pH 和离子强度。缓冲液浓度太低，区带泳动速度加快并增加区带宽度；浓度过高则使区带移动速度减慢，造成区带不易分离。要特别注意环境温度升高时，因缓冲液蒸发而使浓度增大造成的影响。如血清蛋白电泳时可选用 pH 8.6 的巴比妥缓冲液或硼酸缓冲液，氨基酸的分离则可选用 pH 7.2 的磷酸缓冲液等。

（3）操作步骤　点样→电泳→染色→漂洗→透明→测定光密度。选用一定长度和宽度的薄膜条，在其一端处点上样品，在一定电压、电泳条件下进行电泳。染色可根据不同物质使用不同的染色方法，如蛋白质可采用氨基黑 10B 或考马斯亮蓝进行染色。染色漂洗后，可用光密度计扫描来进行定量测定。

四、琼脂糖凝胶电泳

琼脂糖是从琼脂中提取出来的，是由 D-半乳糖和 3,6-脱水-L-半乳糖结合的链状多糖，含硫酸根比琼脂少，因而分离效果明显提高。

琼脂糖电泳具有以下优点：①琼脂糖含液体量大，可达 98%～99%，近似自由电泳，但样品的扩散度比自由电泳小，对蛋白质的吸附极微；②琼脂糖作为支持体有分辨率高、重复性好的特点；③电泳速度快；④透明而不吸收紫外线，可以直接用紫外检测仪作定量测定；⑤区带可染色，样品易回收，有利于制备。缺点是琼脂糖中有较多硫酸根，电渗作用大。

琼脂糖凝胶电泳也常用于分离、鉴定核酸，如 DNA 鉴定、DNA 限制性内切酶图谱制作等，为 DNA 分子及其片段的分子量测定和 DNA 分子构象的分析提供了重要手段。由于这种方法具有操作方便、设备简单、需样品少、分辨能力高的优点，已成为基因工程研究中的常用方法之一。

琼脂糖凝胶电泳对核酸的分离作用主要依据它们的分子量及分子构型，同时与凝胶的浓度也有密切关系。

(一) 核酸分子大小与琼脂糖浓度的关系

1. DNA 分子的大小

在凝胶中，较小的 DNA 片段迁移比较大的片段快。DNA 片段迁移距离（迁移率）与其分子量的对数成反比，见图 1-2。因此，通过已知大小的标准物移动的距离与未知片段的移动距离进行比较，便可测出未知片段的大小。但是当 DNA 分子大小超过 20kb 时，琼脂糖凝胶就很难将它们分开，此时电泳的迁移率不再依赖于分子大小。因此，应用琼脂糖凝胶电泳分离 DNA 时，分子大小不宜超过此值。

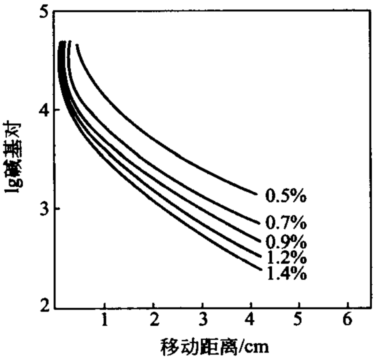


图 1-2　移动距离与碱基对的相应关系
0.5×TBE 缓冲液，含 0.5μg/ml 溴化乙锭；电泳条件：1V/cm，16h。

2. 琼脂糖的浓度

一定大小的 DNA 片段在不同浓度的琼脂糖凝胶中，电泳迁移率不相同。不同浓度的琼脂糖凝胶适宜分离的 DNA 片段大小范围详见表 1-3。要有效地分离大小不同的 DNA 片段，主要是选用适当的琼脂糖凝胶浓度。

表 1-3　琼脂糖凝胶浓度与分辨 DNA 大小范围的关系

琼脂糖凝胶浓度/%	可分辨的线性 DNA 大小范围/kb	琼脂糖凝胶浓度/%	可分辨的线性 DNA 大小范围/kb
0.3	60～5	1.2	6～0.4
0.6	20～1	1.5	4～0.2
0.7	10～0.8	2.0	3～0.1
0.9	7～0.5		

(二) 核酸构型与琼脂糖凝胶电泳分离的关系

不同构型的 DNA 在琼脂糖凝胶中的电泳速度差别较大。在分子量相同的情况下，不同构型的 DNA 的移动速度次序如下：共价闭环 DNA（简称 cccDNA）>线性 DNA>开环的双链环状 DNA。

（三）琼脂糖凝胶电泳基本方法

1. 凝胶电泳类型

用于分离核酸的琼脂糖凝胶电泳也可分为垂直型及水平型（平板型）两种。水平型电泳时，凝胶板完全浸泡在电泳缓冲液下1～2mm，故又称为潜水式电泳。目前更多用的是水平型电泳，因为它制胶和加样比较方便，电泳槽简单，易于制作，又可以根据需要制备不同规格的凝胶板，节约凝胶，因而受到人们的欢迎。

2. 缓冲液系统

DNA的电泳迁移率受到电泳缓冲液的成分和离子强度的影响，当离子强度低时，DNA迁移快；相反，离子强度高时，DNA迁移慢，但由于高离子强度电导很高，导致大量热量产生，严重时，会造成胶熔化和DNA的变性。

常用电泳缓冲液的配制见表1-4。电泳缓冲液一般都配制成浓的贮备液，临用时稀释到所需倍数。

表1-4 常用的琼脂糖凝胶电泳缓冲液

缓冲液	工作溶液	贮存液(1000ml)
Tris-乙酸 (TAE)	1×:0.04mol/L Tris-乙酸 0.001mol/L EDTA	50×:242g Tris 57.1ml 冰乙酸 100ml 0.5mol/L EDTA(pH8.0)
Tris-磷酸 (TPE)	1×:0.09mol/L Tris-磷酸 0.002mol/L EDTA	10×:108g Tris 15.5ml 85%磷酸(1.679g/ml) 40ml 0.5mol/L EDTA(pH8.0)
Tris-硼酸 (TBE)	0.5×:0.045mol/L Tris-硼酸 0.001mol/L EDTA	5×:54g Tris 27.5g 硼酸 20ml 0.5mol/L EDTA(pH8.0)

3. 样品配制与加样

DNA样品用适量Tris-EDTA缓冲液溶解，缓冲液内含有0.25%溴酚蓝或其他指示染料，染料中含10%～15%蔗糖或10%～15%甘油，以增加加样液的密度，使样品集中。

4. 电泳

琼脂糖凝胶分离大分子DNA实验条件的研究结果表明：在低浓度、低电压下，分离效果较好。在低电压条件下，线性DNA分子电泳迁移率与所用的电压呈正比。但是，在电场强度增加时，分子量高的DNA片段迁移率的增加是有差别的。因此，随着电压的增高，电泳分辨率反而下降，分子量与迁移率之间就可能偏离线性关系。为了获得电泳分离DNA片段的最大分辨率，电场强度不宜高于5V/cm。

5. 染色

常用荧光染料溴化乙锭（EB）进行染色，以观察琼脂糖凝胶内的DNA条带。

五、常规聚丙烯酰胺凝胶电泳

聚丙烯酰胺凝胶是由单体丙烯酰胺（acrylamide，简称Acr）和交联剂[又称为共聚体的N,N'-亚甲基双丙烯酰胺（methylene-bisacrylamide，简称Bis）]在加速剂和催化剂的作用下聚合交联成的三维网状结构凝胶，以此凝胶为支持物的电泳称为聚丙烯酰胺凝胶电泳，简称PAGE。与其他凝胶相比，聚丙烯酰胺凝胶有下列优点：①在一定浓度时，凝胶透明，有弹性，机械性能好；②化学性能稳定，与被分离物不起化学反应；③对pH和温度变化较

稳定；④几乎无电渗作用，只要 Acr 纯度高，操作条件一致，则样品分离重复性好；⑤样品不易扩散，且用量少，其灵敏度可达 10^{-6}g；⑥凝胶孔径可调节，根据被分离物的分子量选择合适的浓度，通过改变单体及交联剂的浓度调节凝胶的孔径；⑦分辨率高，尤其在不连续凝胶电泳中，集浓缩、分子筛和电荷效应为一体，因而较醋酸纤维素薄膜电泳、琼脂糖凝胶电泳等有更高的分辨率。

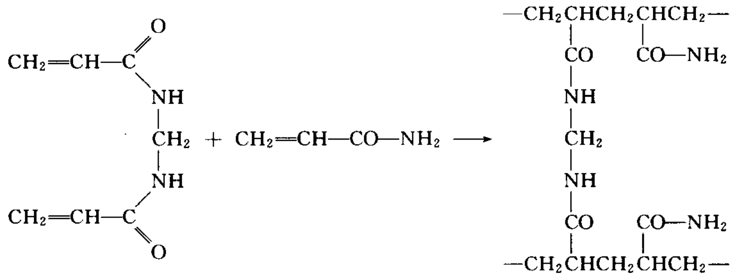
PAGE 常用于蛋白质、酶、核酸等生物分子的分离、定性、定量及少量的制备等。

除了常规聚丙烯酰胺圆盘及垂直板电泳，还有聚丙烯酰胺梯度凝胶电泳、SDS-聚丙烯酰胺凝胶电泳、等电聚焦电泳及双向凝胶电泳等技术，这些技术在许多方面有共同之处，但又有各自的特点，将在后面分别介绍。

（一）聚丙烯酰胺凝胶聚合原理及相关特性

1. 聚合反应

聚丙烯酰胺凝胶（polyacrylamide gel，PAG）是由丙烯酰胺单体（Acr）和交联剂亚甲基双丙烯酰胺（Bis）在催化剂过硫酸铵或核黄素作用下聚合交联而成的三维网状结构凝胶，其聚合反应如图 1-3 所示。



N,*N*′-亚甲基双丙烯酰胺　　丙烯酰胺　　聚丙烯酰胺凝胶

图 1-3　丙烯酰胺和亚甲基双丙烯酰胺的聚合反应

聚合反应时常用的催化剂有过硫酸铵及核黄素两个系统。在水溶液中，过硫酸铵离子 $S_2O_8^{2-}$ 可形成游离基 SO_4^{2-}，它能使丙烯酰胺单体的双键打开，形成游离基丙烯酰胺，后者和 Bis 单体作用，能聚合成凝胶。催化反应需要在碱性条件下进行，如用 7％的丙烯酰胺，在 pH8.3 时，30min 就能聚合完毕。为避免溶液中有氧气而妨碍聚合，在反应前可将溶液抽气除氧。核黄素在光照下部分分解并被还原成无色型核黄素；但在有氧的条件下此无色型核黄素又被氧化成为带有游离基的核黄素，后者也能使丙烯酰胺和亚甲基双丙烯酰胺聚合成凝胶。为加速聚合，还加入四甲基乙二胺（TEMED）作为加速剂，促进聚合作用。

聚丙烯酰胺凝胶因富含酰胺基，使凝胶具有稳定的亲水性。它在水中无电离基团，不带电荷，几乎没有吸附及电渗作用，是一种比较理想的电泳支持物。

2. 凝胶孔径的可调性及其有关性质

（1）凝胶性能与总浓度及交联度的关系　凝胶的孔径、机械性能、弹性、透明度、黏度和聚合程度取决于凝胶总浓度和 Acr 与 Bis 之比。通常用 *T*（％）表示总浓度，即 100ml 凝胶溶液中含有 Acr 及 Bis 的总质量（g）。Acr 和 Bis 的比例常用交联度 *C*（％）表示，即交联剂 Bis 占单体 Acr 与 Bis 总量的比例。要想将蛋白质或核酸之类的大分子混合物很好地分离，并在凝胶上形成明显的区带，选择一定孔径的凝胶是个关键。常用的标准凝胶是指浓度为 7.5％的凝胶，大多数生物体内的蛋白质在此凝胶中电泳，能获得满意的结果。

(2) 凝胶浓度与被分离物分子量的关系　由于凝胶浓度不同，平均孔径不同，能通过的可移动颗粒的分子量也不同，其大致范围如表 1-5。在操作时，可根据被分离物质的分子量大小选择所需凝胶的浓度范围。也可先选用 7.5%凝胶（标准胶），因为生物体内大多数蛋白质在此范围内电泳均可获得较满意的结果。

表 1-5　被分离物质的分子量范围与聚丙烯酰胺凝胶浓度的关系

被分离物质的分子量范围		适用的凝胶浓度/%	被分离物质的分子量范围		适用的凝胶浓度/%
蛋白质	$<10^4$	20～30	核酸(RNA)	$<10^4$	15～20
	$(1\sim4)\times10^4$	15～20		$10^4\sim10^5$	5～10
	$(1\sim5)\times10^4\sim1\times10^5$	10～15		$10^5\sim2\times10^6$	2～2.6
	1×10^5	5～10			
	$>5\times10^5$	2～5			

（二）聚丙烯酰胺凝胶电泳原理

聚丙烯酰胺凝胶电泳（polyacryamide gel electrophoresis，简称 PAGE）根据其有无浓缩效应，分为连续系统与不连续系统两大类。前者电泳体系中缓冲液 pH 及凝胶浓度相同，带电颗粒在电场作用下，主要靠电荷及分子筛效应分离；后者电泳体系中由于缓冲液离子成分、pH、凝胶浓度及电位梯度的不连续性，带电颗粒在电场中泳动不仅有电荷效应、分子筛效应，还具有浓缩效应，因而其分离条带清晰度及分辨率均较前者为佳。目前常用的电泳方式多为垂直的圆盘及板状两种。前者凝胶是在玻璃管中聚合，样品分离区带染色后呈圆盘状，因而称为圆盘电泳（disc electrophoresis），其装置如图 1-4。后者凝胶是在两块间隔几毫米的平行玻璃板中聚合，故称为板状电泳（slab electrophoresis）。两者电泳原理完全相同。现以 pH 不连续圆盘 PAGE 分离血清蛋白为例，阐明各种效应的原理。

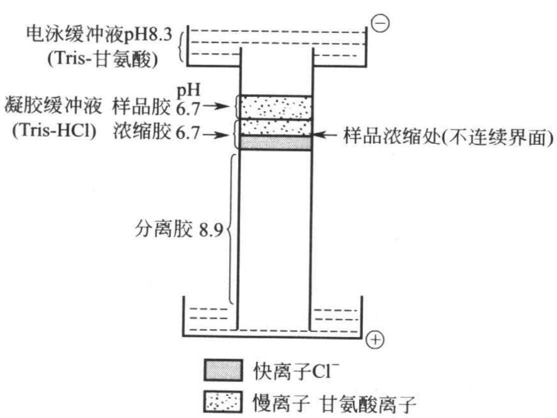


图 1-4　不连续圆盘电泳示意图

不连续体系由电泳缓冲液、样品胶、浓缩胶及分离胶所组成，在直立的玻璃管中（或 2 层玻璃板中）依次排列为上层样品胶、中间浓缩胶、下层分离胶，如图 1-4 和图 1-5 所示。

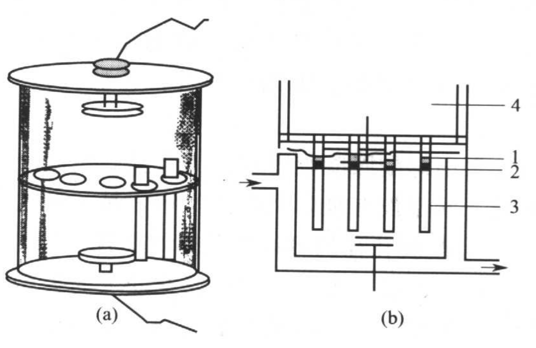


图 1-5　聚丙烯酰胺凝胶圆盘电泳示意图

(a) 电泳系统正面，(b) 电泳系统剖面

1—样品胶，pH6.7；2—浓缩胶，pH6.7；

3—分离胶，pH8.9；4—电泳缓冲液，pH8.3

样品胶是聚合成的大孔胶，$T=3\%$，$C=2\%$，其中含有一定量的样品及 pH6.7 的 Tris-HCl 凝胶缓冲液，其作用是防止对流，促使样品浓缩以免被电泳缓冲液稀释。目前，一般不用样品胶，而是直接在样品液中加入等体积 40%蔗糖，同样具有防止对流及防止样品被稀释的作用。

实际上，浓缩胶是样品胶的延续，凝胶浓度及 pH 与样品胶完全相同，其作用是使样品进入分离胶前，被浓缩成窄的扁盘，从而提高分离效果。

分离胶是聚合成的小孔胶，$T=7.0\%\sim7.5\%$，$C=2.5\%$，凝胶缓冲液为 pH8.9 的 Tris-HCl，大部分血清蛋白质在此 pH 条件下，按各自负电荷量及分子量泳动。此胶主要起分子筛作用。

上、下电泳槽是用聚苯乙烯或二甲基丙烯酸制作的。将带有 3 层凝胶的玻璃管垂直放在电泳槽中，在两个电极槽倒入足够量 pH8.3 的 Tris-甘氨酸电泳缓冲液，接通电源即可进行电泳。在此电泳体系中，有 2 种孔径的凝胶、2 种缓冲体系、3 种 pH，因而形成了凝胶孔径、pH、缓冲液离子成分的不连续性，这是样品浓缩的主要因素。PAGE 具有较高的分辨率，主要是因为在电泳体系中集样品浓缩效应、分子筛效应及电荷效应为一体。

目前，PAGE 连续体系应用也很广，虽然电泳过程中无浓缩效应，但利用分子筛及电荷效应也可使样品得到较好的分离，加之在温和的 pH 条件下进行，不致使蛋白质、酶、核酸等活性物质变性失活，故也常在实验中采用。

聚丙烯酰胺凝胶垂直板电泳是在圆盘电泳的基础上建立的，两者电泳原理完全相同，只是灌胶的方式不同，垂直板电泳的凝胶不是灌在玻璃管中，而是灌在嵌入橡胶框凹槽中长度不同的两块平行玻璃板的间隙内，且间隙可调节，一般有 0.5mm、1.5mm 及 3mm 三种规格的橡胶框，前两种多用于分析鉴定，最后一种常用于制备。垂直板电泳较圆盘电泳有更多的优越性：

① 表面积大而薄，便于通冷却水以降低热效应，条带更清晰；

② 在同一块胶板上，可同时进行 10 个以上样品的电泳，便于在同一条件下比较分析鉴定，还可用于印迹转移电泳及放射自显影；

③ 胶板制作方便，易剥离，样品用量少，分辨率高，既可用于分析，又可用于制备；

④ 胶板薄而透明，电泳染色后可制成干板，便于长期保存与扫描。

血清蛋白在纸电泳或醋酸纤维素薄膜电泳中只能分离出 5～6 条区带，而在上述两种形式的聚丙烯酰电泳中却可分离出数十条区带。

（三）操作方法

1. 制胶（以垂直板不连续电泳为例）

不连续电泳是将浓缩胶加在分离胶上，并且使用不同缓冲系统，所以需要分别灌注浓缩胶和分离胶。表 1-6 为不连续电泳的凝胶配方，贮存液的配制如下。

（1）丙烯酰胺单体贮存液（$T=30\%$，$C=3\%$）　14.55g 丙烯酰胺和 0.45g N,N'-亚甲基双丙烯酰胺，先用 40ml 双蒸水搅拌溶解，直到溶液变成透明，再用双蒸水稀至 50ml，过滤。用棕色瓶可在 4℃保存 1 个月。丙烯酰胺和亚甲基双丙烯酰胺单体及溶液是中枢神经毒物，要小心操作。

（2）浓缩胶缓冲液贮存液（0.5mol/L Tris-HCl，pH6.8）　3.03g Tris 溶解在 40ml 双蒸水中，用 4mol/L 盐酸调 pH 至 6.8，再用双蒸水稀至 50ml，保存在 4℃备用。

（3）分离胶缓冲液贮存液（1.5mol/L Tris-HCl，pH8.9）　18.16g Tris 溶解在 80ml 双蒸水中，用 4mol/L 盐酸调 pH 至 8.9，再用双蒸水稀至 100ml，保存在 4℃备用。

（4）10%过硫酸铵　0.1g 过硫酸铵加 1.0ml 双蒸水。使用前新鲜配制。

（5）电泳缓冲液（0.025mol/L Tris，0.2mol/L 甘氨酸，pH8.3）　15.14g Tris 加 72.07g 甘氨酸，用双蒸水稀释到 5L；可在室温保存 1 个月。

表 1-6 不连续电泳浓缩胶和分离胶的配方

试剂用量	凝胶终浓度 T(C=3%)				
	4%	7.5%	10%	15%	20%
单体贮存液/ml	0.65	3.8	5.0	7.5	10.0
浓缩胶缓冲液贮存液/ml	0.10	—	—	—	—
分离胶缓冲液贮存液/ml	—	0.3	0.3	0.3	0.3
双蒸水/ml	4.25	10.9	9.7	7.2	4.7
10%过硫酸铵/μl	15	15	15	15	15
TEMED/μl	7	7	7	7	7
总体积/ml	5	15	15	15	15

注：1. 在加过硫酸铵和 TEMED 以前，溶液最好抽气。
2. 过硫酸铵和 TEMED 的量应根据室温和聚合情况而定，控制在 30min 左右凝胶聚合。过硫酸铵必须新鲜配制。
3. 欲配制不同终浓度、不同厚度或不同大小的凝胶，可根据表中的体积按比例变化。
4. 凝胶聚合后，应在 4℃保存 12h 再使用，以使凝胶充分聚合，改善电泳时的分辨率。
5. 表中 T 为 7.5%、10%、15%、20%（C=3%）的分离胶可作为连续电泳用的分离胶。

（6）样品缓冲液（0.1mol/L Tris-HCl，pH6.8） 2ml 浓缩胶缓冲液贮存液加入 1ml 87%甘油和 0.1mg 溴酚蓝，用双蒸水稀释至 10ml；可在−20℃保存 6 个月。

按配方在模具中灌注分离胶后，小心地在分离胶的表面加一层水，封住胶面，以促使聚合并使凝胶表面平直。凝胶放置约 30min 左右聚合，这时可以重新看到一个界面，表示凝胶已经聚合。凝胶充分聚合后再吸掉上层水分，用浓缩胶缓冲液贮存液淋洗凝胶，然后灌注浓缩胶，并插入与模具大小相同、与凝胶厚度相当的梳子。为防止气泡陷入，梳子应倾斜插入。然后让模具在 30～40℃静止放置，浓缩胶也应该在 30min 左右聚合。

2. 样品的准备及加样

常规 PAGE 的样品一般不需作特殊处理，如果样品溶液带有较高浓度的盐，则应先用透析或凝胶过滤柱脱盐。

在样品缓冲液加入样品，如样品溶解性不好，可加入少量 TrtionX-100 及 NP-40 助溶。准备好的样品在 Eppendorf 管中 10000r/min 离心 3～5min，取上清加样，以免电泳时产生拖尾。

加样量随凝胶板的厚度及加样孔的宽度、染色方法而异，银染所需的样品量仅为考马斯亮蓝的 1/100～1/20。

加样前轻轻把梳子拔出，用电泳缓冲液淋洗加样孔，吸出，再加适量的电泳缓冲液，然后用微量移液器小心加入样品，不要带入气泡。

3. 电泳

小心地在上槽中加入电泳缓冲液，切不可冲散加样孔中的样品。打开冷却循环系统，连接电源。常采用较低的起始电压开始电泳，以利于样品进入凝胶，待样品全部进胶后再升高到正常值。一般垂直电泳需要 3～6h，待溴酚蓝前沿到达凝胶板底部（阳极）时，关闭电源，关掉冷却系统，取出凝胶，准备染色。

4. 检测

蛋白质样品电泳后的检测方法主要是进行染色。目前以考马斯亮蓝染色最为常用。对于样品量少、需要高灵敏度染色的电泳凝胶，则采用银染色。

六、SDS-聚丙烯酰胺凝胶电泳原理

各种蛋白质因所带的净电荷、分子量大小和形状不同而有不同的迁移率。如果在电泳体

系加入十二烷基硫酸钠（sodium dodecyl sulfate，SDS），则电泳迁移率主要依赖于分子量，与所带的净电荷和形状无关，这种电泳方法称为 SDS-PAGE，主要用于测定蛋白质亚基分子量。

SDS 是一种阴离子去污剂，作为变性剂和助溶性试剂，它能断裂分子内和分子间的氢键，使分子去折叠，破坏蛋白质分子的二级和三级结构。强还原剂，例如巯基乙醇（β-mercapto ethanol）和二硫苏糖醇（dithiothreitol，DTT）则能使半胱氨酸残基之间的二硫键断裂。在样品和凝胶中加入 SDS 和还原剂后，分子被解聚成组成它们的多肽链。解聚后的氨基酸侧链与 SDS 充分结合，形成带负电荷的蛋白质-SDS 复合物。由于 SDS 带有大量负电荷，当其与蛋白质结合时，所带的负电荷大大超过了天然蛋白质原有的负电荷，因而消除或掩盖了不同种类蛋白质间原有电荷的差异，使其均带有相同密度的负电荷。蛋白质-SDS 复合物在水溶液中的形状像一个长椭圆棒，椭圆棒的短轴对不同的蛋白质亚基-SDS 复合物基本上是相同的，约为 1.8nm。但长轴的长度则与亚基分子量的大小成正比。因此这种复合物在 SDS-聚丙烯酰胺凝胶系统中的电泳迁移率不再受蛋白质原有电荷和形状的影响，而主要取决于椭圆棒的长轴长度，即蛋白质或亚基分子量的大小。当蛋白质的分子量在 15kDa 到 200kDa 之间时，电泳迁移率与分子量的对数呈线性关系，这个规律对大部分蛋白质都是适用的。

SDS-PAGE 成功的关键之一是电泳过程中（特别是样品制备过程中）蛋白质与 SDS 的结合程度。影响它们结合的因素主要有三个。

（1）溶液中 SDS 单体的浓度　SDS 在水溶液中是以单体和 SDS-多肽胶束的混合形式存在的，能与蛋白质分子结合的是单体。单体的浓度与 SDS 总浓度、温度和离子强度有关。在一定温度和离子强度下，当 SDS 总浓度增加到某一定值时，溶液中的单体浓度不再随 SDS 总浓度的增加而升高。当单体浓度大于 1mmol/L 时，大多数蛋白质与 SDS 结合的质量比为 1∶1.4。如果单体浓度降到 0.5mmol/L 以下时，二者的结合比仅为 1∶0.4，这样就不能消除蛋白质分子原有的电荷差别，也就不能进行分子量测定。

（2）样品缓冲液的离子强度　因为 SDS 结合到蛋白质分子上的量仅决定于平衡时 SDS 的单体浓度，不是总浓度，而只有在低离子强度的溶液中，SDS 单体才具有较高的平衡浓度，所以 SDS-PAGE 的样品缓冲液离子强度较低，常为 10～100mmol/L。

（3）二硫键是否完全被还原　只有二硫键被彻底还原后，蛋白质分子才能被解聚，SDS 才能定量地结合到亚基上，使蛋白质的相对迁移率和分子量的对数呈线性关系。

七、聚丙烯酰胺梯度凝胶电泳

用连续 SDS-PAGE 测定蛋白质分子量，由于 SDS 及巯基乙醇的作用，天然蛋白质解离为亚基或肽链，因此测得的分子量不是天然蛋白质的分子量，要确定其真正的分子量还需配合其他方法验证。为弥补这一缺陷，Margolis 和 Slater 等以聚丙烯酰胺为支持物，制备成孔径梯度（pore gradient，PG）或称为梯度凝胶，进行 PAGE（简称 PG-PAGE）分离和鉴定各种蛋白质组分，并用来测定蛋白质分子量。实验证实梯度凝胶有更好的分辨率。

线性梯度凝胶制备不同于均一浓度凝胶制备，应预先配制低浓度胶（2%～4%）贮存液置贮液瓶中；高浓度胶（16%～30%）贮存液置混合瓶中（两者体积比为 1∶1），在梯度混合器及蠕动泵的协助下，从下至上灌胶，如图 1-6 所示。

凝胶聚合后，则从下到上形成从浓至稀依次排列的线性梯度凝胶。在 pH 大于蛋白质

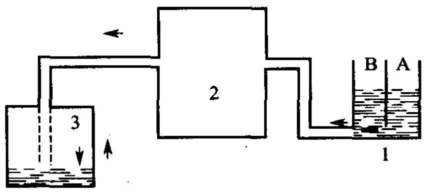


图 1-6 梯度凝胶电泳凝胶装置示意图
1—混合器 A 为贮液瓶，B 为混合瓶；
2—蠕动泵；3—凝胶模。箭头表示
液体流动的方向

p*I* 的缓冲体系中电泳时，蛋白质样品从负极向正极移动，也就是说从上向下，向着凝胶浓度增加（孔径逐渐减小）的方向移动，随着电泳的连续进行，蛋白质颗粒的迁移由于孔径渐小而使阻力愈来愈大。在开始时，蛋白质在凝胶中的迁移速度主要受两个因素影响：一是蛋白质本身的电荷密度，电荷密度愈高，迁移率愈快；二是蛋白质本身的大小，分子量愈大，迁移率速率愈慢。当蛋白质迁移到所受阻力最大时，则完全停止前进。此时，低电荷密度的蛋白质将“赶上”与它大小相似、但具有较高电荷密度的蛋白质。因此，在梯度凝胶电泳中，蛋白质的最终迁移位置仅取决于其本身分子大小，而与蛋白质本身的电荷密度无关。梯度电泳原理可用图 1-7 表示。

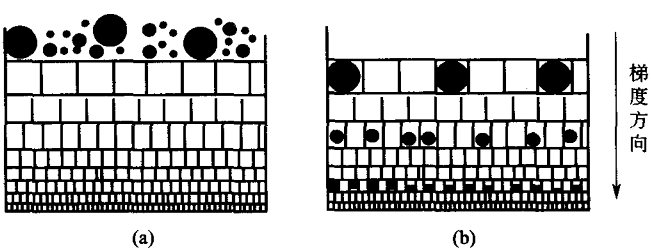


图 1-7 梯度凝胶电泳示意图
（a）电泳开始前；（b）电泳结束时

图 1-7 中方格代表凝胶孔径，自上而下，孔径逐渐变小，形成梯度。图中圆球分别代表大、中、小三种不同分子量的蛋白质。（a）代表电泳开始前分子的状况；（b）代表经过一定时间电泳后，所有大小不同的分子均进入梯度凝胶孔径中，大、中、小分子分别滞留在与分子大小相当的凝胶孔径中，不再前进。将标准分子量蛋白质与未知样品同时电泳，染色后，根据标准蛋白质的相对迁移率与其分子量的对数作图（标准曲线），即可从未知样品相对迁移率查出其分子量。PG-PAGE 与均一的其他类型 PAGE 比较有下列优点。

① 由于梯度凝胶孔径的不连续性，可使样品中各组分充分浓缩，即使样品很稀，在电泳过程中，分二、三次加样，最终样品均滞留于其相应的凝胶孔径中而得到分离。

② 可提供更清晰的蛋白质区带，用于蛋白质纯度的鉴定。

③ 可在一个凝胶板上同时测定数个分子量相差很大的蛋白质。例如 4%～30%PG-PAGE 可分辨分子量为 50000～200000 之间的各种蛋白质。

④ 可直接测定天然蛋白质分子量，蛋白质不被解离为亚基。因此，本法可作为 SDS-PAGE 测定蛋白质分子量的补充。

尽管本法有上述优点，但主要适用于测定球状蛋白质分子量，对纤维状蛋白分子量的测定误差较大。另外，电泳时要求足够高的电压和足够长的时间（一般不低于 2000V·h），否则将达不到预期的效果。

八、聚丙烯酰胺凝胶等电聚焦电泳

蛋白质在 pH 梯度聚丙烯酰胺凝胶中电泳时，因为各种蛋白质 p*I* 不同，当迁移至其 p*I*

相同的pH处则不再泳动，而浓缩成狭窄的区带，这种分离蛋白质的方法称为聚丙烯酰胺凝胶等电聚焦电泳（isoelectric focusing-PAGE，简称IEF-PAGE）。

（一）等电聚焦的原理

在常规PAGE中，分离通常是在pH恒定的缓冲系统中进行，如图1-8（左）所示。假定缓冲液的pH是9，根据净电荷-pH曲线（滴定曲线），蛋白质A和B将分别有净电荷-3和-1。当这些带负电的蛋白质被放在靠近电泳系统的阴极时，它们将在电场的影响下向阳极迁移。假定蛋白质A和B的大小和形状相同，带负电荷多的蛋白质A将受到较大的电场力而跑到蛋白质B的前面，并与蛋白质B分开。由于是在恒定缓冲系统中分离，很易扩散，为此样品必须被加成一个窄带，否则在电泳过程中蛋白带会变宽而影响分离。

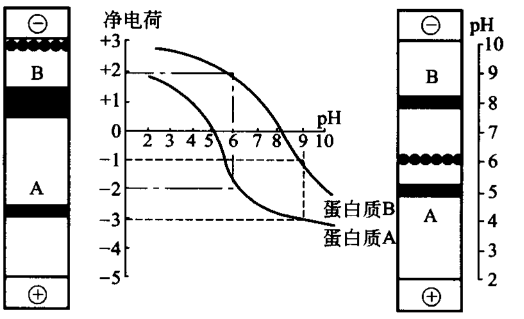


图1-8　常规PAGE（左）和IEF-PAGE（右）原理的区别

在等电聚焦中，分离是在连续、稳定的pH梯度中进行的。如果把蛋白质A和B的混合物放在pH3.5～10的梯度中pH6的位置（图1-8右），根据滴定曲线，蛋白质A将有净电荷－2，蛋白质B将有净电荷＋2。加上电场后，它们分别向不同的方向迁移。A向阳极，B向阴极。在迁移过程中，蛋白质A将进入越来越低的pH范围，并将逐渐损失它的负电荷，当它达到pH5的梯度位置时，它的净电荷为零，即停止迁移，这就是它的p*I*位置。与此同时，蛋白质B将向越来越高的pH范围迁移，直到到达它的净电荷是零的pH梯度位置，这就是蛋白质B的等电点位置。

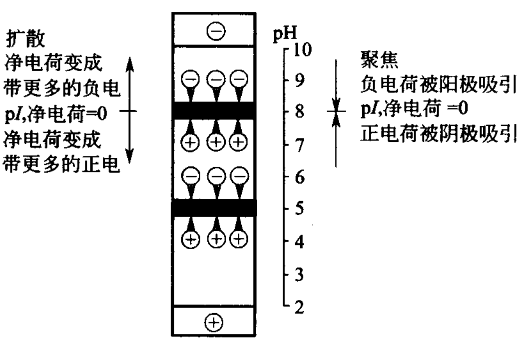


图1-9　等电聚焦的“聚焦效应”

在等电聚焦中，分离仅仅决定于蛋白质的等电点，一旦蛋白质到达它的等电点位置，将失去净电荷，就不能进一步迁移。如果蛋白带向阴极扩散，则将进入高pH范围而带负电，阳极就将吸引它回去，直到它回到净电荷是零的位置。同样如果它向阳极扩散，则将带上正电，阴极将吸引它回去，直到它回到净电荷是零的位置。因此蛋白质只能在它的等电点位置被聚焦成一条窄而稳定的带，见图1-9。所以，等电聚焦的分辨率大大高于常规聚丙烯酰胺凝胶电泳。这种“聚焦效应”或称“浓缩效应”是等电聚焦最大的优点。是高分辨率的保证。

（二）pH梯度的形成

等电聚焦技术的关键在于凝胶中pH梯度的建立，根据建立pH梯度原理的不同，又可分为载体两性电解质pH梯度和固相pH梯度。

最常用的载体两性电解质是“Ampholine”，有不同的pH范围可供选用，它的pH梯度是在电场中通过两性缓冲离子在电泳的过程中建立起来的。电泳前，载体两性电解质溶液的pH大约是该溶液pH范围的平均值。所有的载体两性电解质分子都荷电，只是在溶液中荷正电和荷负电的基团数目是相等的，所以总的净电荷将是零，见图1-10(a)。当加上电场时，

载体两性电解质分子将向阴极或阳极迁移，带有最低等电点的分子（荷最多的负电）将最快地向阳极迁移。当它达到净电荷是零的位置时才停止。这个位置靠近阳极，由于它有高的缓冲能力，使环境溶液的 pH 等于分子本身的 p*I*。其次一些低 p*I* 的载体两性电解质分子也将向阳极移动，直到它的净电荷被减少到零才停止。同样地，由于它的高的缓冲能力，也将使环境溶液的 pH 等于它本身的 p*I* 见图 1-10(b)。依此类推，所有的载体两性电解质分子用增加 p*I* 级数的办法将分别在阳、阴极之间到达它们自己的位置而形成一个 pH 梯度，见图 1-10(c)。Ampholine 的 pH 梯度开始形成在正、负极两端，然后移向中间。在 12cm 凝胶距离中，使用 30W 功率，在 1h 后 pH 梯度完全形成，在 2h 后，pH 梯度只有很小的改变。

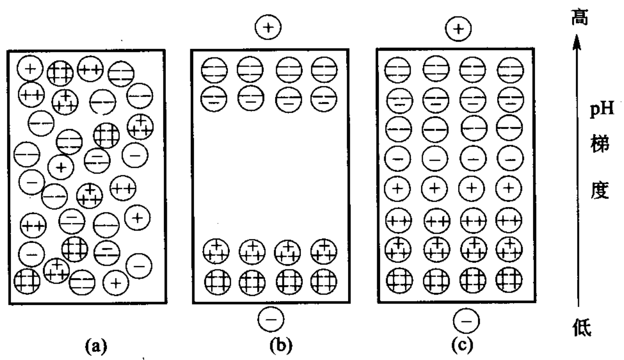


图 1-10 载体两性电解质在电场中形成 pH 梯度的模式图

固相 pH 梯度等电聚焦（immobilized pH gradients isoelectric focusing，IPGIEF）是 20 世纪 80 年代初建立起来的一种新型等电聚焦技术。Immobiline 是丙烯酰胺的衍生物，利用一系列具有弱酸或弱碱性质的丙烯酰胺衍生物滴定时，在滴定终点附近形成 pH 梯度并参与丙烯酰胺的共价聚合，从而形成固定的、不随环境电场等条件变化的 pH 梯度。该方法具有比传统载体两性电解质等电聚焦更高的分辨能力、更大的上样量，pH 分辨率可达 0.001，可分析和制备等电点极其相近的蛋白质和多肽等，主要用于双向电泳中的第一向等电聚焦电泳。固相 pH 梯度凝胶制备技术比较复杂，一般均使用商品 IPG 干胶条。

九、双向凝胶电泳

双向凝胶电泳又常被称为二维凝胶电泳，第一向是以蛋白质电荷差异为分离机理的等电聚焦（载体两性电解质 pH 梯度或固相 pH 梯度）电泳，第二向是以蛋白质分子量差异为分离机理的 SDS-PAGE。样品分别经过电荷和分子量的两次分离后，分离的结果不是条带，而是斑点，从左到右是 p*I* 的增加，从下到上是分子量的增加。它的分辨率极高，可达 1000～3000 点，这是目前所有电泳技术中分辨率最高、信息最多的技术，是蛋白质组研究的关键技术，详情请见第十二章“蛋白质组学技术”。双向凝胶电泳的基本步骤如下。

1. 样品制备

由于双向电泳要对样品从两个方向上依次进行 IEF-PAGE 和 SDS-PAGE，这样对要分析的样品就比单独进行 IEF-PAGE 或 SDS-PAGE 有更高的要求。样品加样量的多少、溶解的好坏、还原是否完全等因素都将很大程度地影响到分离的结果。样品制备大体上包括蛋白质的溶解、变性及还原、去除非蛋白质杂质等。样品溶解得不好会减少分离到的蛋白质数

量，同时会造成等电聚焦时某些蛋白质的沉淀，从而减少转移到第二向的蛋白质数量。

2. 第一向等电聚焦电泳

与单纯的 IEF-PAGE 基本相同，第一向等电聚焦可采用传统的 Ampholine 的 pH 梯度管状凝胶（直径 1～2mm）来进行。但为了满足蛋白质组研究所要求的高度重复性，现在第一向均采用具有固相 pH 梯度的 IPG 干胶条及配套的仪器，使等电聚焦基本实现自动化，并且重复性良好。目前通常采用的胶条厚 0.5mm、宽 3mm，长度有 7cm、11cm、13cm、18cm 和 24cm 等规格；pH 范围有宽、窄两类，宽范围通常为 3～10，还有各种窄 pH 范围的胶条，如 4～7、6～10 等。窄 pH 范围通常有更好的分辨率，上样量也可以加大，一般用于细胞蛋白质更精确的分离和分析。

3. 平衡及两向间的转移

第一向 IEF-PAGE 完成后，取出凝胶，用含有 SDS、在还原条件下的 pH6.8 Tris 缓冲液平衡，目的是使蛋白质分子与 SDS 充分相互作用；平衡时间不可太短，柱状胶约需 30～40min，薄层胶（0.5～1mm）约需 5～10min。

平衡后的等电聚焦凝胶与 SDS 凝胶的良好接触是双向电泳结果的保证。将等电聚焦凝胶贴于 SDS 凝胶上面，避免夹有气泡，防止胶条被拉长，管状凝胶的转移可采用琼脂糖密封。为了以后分子量的测定，在 SDS 凝胶的一端可加蛋白分子量标准。

4. 第二向 SDS-PAGE

与单纯的 SDS-PAGE 基本相同。为了蛋白质得到更好的分离，使用的胶也越来越大（25cm×25cm），以得到更清晰的分离图谱。

5. 蛋白质的检测及图谱数字化分析

电泳后，凝胶上蛋白质可采用考马斯亮蓝染色、银染、荧光标记、放射自显影等多种检测方法。双向电泳分辨率高，分离的斑点很难用肉眼来比较和辨别。蛋白染色后的图谱需要经过图像扫描，计算机数字化处理，确定每个蛋白质斑点的等电点和分子量，提供蛋白质鉴定的初步信息，建立起蛋白质组数据库。一旦某些蛋白质斑点经过分析鉴定，确认其为感兴趣的蛋白质，即可将这些斑点取出，酶解后再进行高效液相色谱-质谱（HPLC-MS）分析，进一步分析该蛋白质分子的分子结构和组成。

十、染色方法

经醋酸纤维素薄膜电泳、琼脂糖凝胶电泳、聚丙烯酰胺凝胶电泳分离的各种生物分子需用染色法使其在支持物相应位置上显示出谱带，从而检测其纯度、含量及生物活性。蛋白质、核酸及酶等均有不同的染色方法，现分别介绍如下。

（一）蛋白质染色

蛋白质染色液种类繁多，染色原理不同，灵敏度各异，使用时可根据需要加以选择。常用的染色液有以下几种。

1. 氨基黑 10B（amino black 10B）

氨基黑 10B 分子式为 $C_{22}H_{13}N_6S_3Na_3$，$M_r=715$，$\lambda_{max}=620\sim630$nm，是酸性染料，其磺酸基与蛋白质反应形成复合盐。氨基黑 10B 是最常用的蛋白质染料之一，但对 SDS-蛋白质染色效果不好。

2. 考马斯亮蓝 R-250（简称 CBB R_{250}）

考马斯亮蓝 R-250 的分子式为 $C_{14}H_{44}O_7H_3S_2Na$，$M_r=824$，$\lambda_{max}=560\sim590$nm，染色

灵敏度比氨基黑高5倍。该染料是通过范德华力与蛋白质结合，尤其适用于SDS电泳微量蛋白质染色，但蛋白质浓度超过一定范围时，对高浓度蛋白质染色不符合比尔定律，作定量分析时要注意这点。

3. 考马斯亮蓝G-250（简称$CBBG_{250}$）

考马斯亮蓝G-250比CBB R_{250}多二个甲基，$M_r=854$，$\lambda_{max}=590\sim610nm$，染色灵敏度不如R-250，但比氨基黑高3倍。其优点是在三氯乙酸中不溶而成胶体，能选择地使蛋白质染色而几乎无本底色，所以常用于重复性好和稳定的染色，适于作定量分析。

4. 荧光染料

（1）丹磺酰氯（2,5-二甲氨基萘磺酰氯，dansyl chloride，简称DNS-Cl） 在碱性条件下与氨基酸、肽、蛋白质的末端氨基发生反应，使它们获得荧光性质，可在波长320nm或280nm的紫外灯下观察染色后的各区带或斑点。蛋白质与肽经丹磺酰化后并不影响电泳迁移率，因此少量丹磺酰化的样品还可用作无色蛋白质分离的标记物。

（2）荧光胺（fluorescamine，又称fluram） 其作用与丹磺酰氯相似。由于自身及分解产物均不显示荧光，因此染色后也没有荧光背景。但由于引进了负电荷，因而引起了电泳迁移率的改变，但在SDS存在下这种电荷效应可忽略。

5. 银染色法

此法是Switzer R. C. 和Merril C. R. 首先提出的，它较CBB R_{250}灵敏100倍，染色机制是将蛋白带上的硝酸银（银离子）还原成金属银，以使银颗粒沉淀在蛋白带上。其灵敏度很高，牛血清蛋白为$4\times10^{-5}\mu g/mm^2$，清蛋白为$8\times10^{-5}\mu g/mm^2$。银染对水和试剂的要求较高，必须使用无离子双蒸水，甲醛和戊二醛必须是新鲜试剂，无聚合，其余试剂也应该用分析纯。染色在日光下进行，并需不断振摇。

银染色的第一步是固定。固定有两个目的：一是将蛋白固定在凝胶中或至少是防止蛋白在凝胶中扩散；第二个目的是去除干扰染色的物质，如去污剂、还原试剂和缓冲液的一些组分（如甘氨酸）。固定液可以是甲醛、乙醇（甲醇）和乙酸或三氯乙酸。固定后的凝胶才能由银颗粒显色。对SDS-PAGE的凝胶进行银染色，一定要在染色前除去SDS，所以通常要采用几次漂洗的程序。

（二）核酸染色

核酸染色法一般可将凝胶先用三氯乙酸、甲酸-乙酸混合液、氯化高汞、乙酸、乙酸镧等固定，或者将有关染料与上述溶液配在一起，同时固定与染色。有的染色液同时染DNA及RNA，如stains-all、溴化乙锭、焙花青-铬矾法等，也有RNA、DNA各自特殊的染色法。

1. RNA染色法

（1）荧光染料溴化乙锭（ethidium bromide，简称EB） EB能与核酸分子中的碱基结合而使核酸在紫外光（253nm）照射下显示出荧光，因此可用于观察琼脂糖电泳中的RNA、DNA带。EB与超螺旋DNA结合能力小于双链开环DNA，而双链开环DNA与EB结合能力又小于线性DNA。用EB染色方法具有操作简单、灵敏度高的优点，对1ng RNA、DNA均可显色。由于EB当料是一种强烈的诱变剂，因此操作时应注意防护，戴上聚乙烯手套。

（2）焦宁Y（pyronine Y） 此染色料对RNA染色效果好，灵敏度高，脱色后凝胶本底颜色浅而RNA色带稳定，抗光且不易退色。焦宁G（pyronine G）也可用于RNA染色。

2. DNA 染色法

除了用 EB 染色外，还有以下几种方法。

(1) 甲基绿 (methyl grenn) 一般将 0.25%甲基绿溶于 0.2mol/L pH4.1 的乙酸缓冲液中，用氯仿抽提至无紫色，将含 DNA 的凝胶浸入，室温下染色 1h 即可显色，此法适用于检测天然 DNA。

(2) 二苯胺 (diphenylamine) DNA 中的 2-脱氧核糖在酸性环境中与二苯胺试剂染色 1h，再在沸水浴中加热 10min 即可显示蓝色区带。此法可区别 DNA 和 RNA。

第三节 色 谱 技 术

色谱技术 (chromatography)，由于翻译和习惯的原因，又常称为层析法或色层法。近百年来，色谱技术不断发展，形式多种多样，相继出现了液相色谱、薄层色谱、气相色谱、离子交换色谱、凝胶色谱、亲和色谱、反相色谱、正相色谱、高效液相色谱，以及色谱与电泳结合的产物——毛细管电色谱。几乎每一种色谱法都已发展成为一门独立的分析技术，在生物大分子领域内得到了广泛的应用。

一、色谱技术的概念

色谱技术是利用混合物中各组分的理化性质的差异（吸附力、溶解度、分子形状和大小、分子极性、分子亲和力等)，各组分在两个相中具有不同的分配系数，其中一相叫作固定相 (stationary phase)，另一相流过此固定相，叫作流动相 (mobile phase)，当溶质在两相间做相对移动时，各物质在两相间进行多次分配，从而使得结构上具有微小差异的各组分得到分离，再配合相应的光学、电学、电化学或其他检测手段，对各组分进行分析。

色谱技术是一种分离、分析的方法。它是一种极好的分离纯化的方法，能把混合物中的各个组分纯化分离开来，还能进行精确的定量，并具有一定的定性分析能力。在色谱中，通常是根据色谱峰的面积或高度进行定量分析，根据色谱峰的位置（保留值）来进行定性分析。

现代色谱技术具有极高的分离效率和灵敏度。高效液相色谱常规分析的灵敏度可达微克至纳克，气相色谱常规分析的灵敏度可达 10^{-11}g，毛细管电泳常规分析的灵敏度可达 10^{-12}g。色谱技术广泛地应用于有机物、无机物、低分子或高分子化合物以及具有生物活性的生物大分子的分离和分析。

二、色谱法的分类

色谱法的种类繁多，分类的方法也多种多样，一般可按两相所处的状态、色谱的分离机制这二种方式进行分类。

(一) 按两相所处的状态分类

色谱法中总是要有两相，流动相可以是液体也可以是气体。按流动相的状态不同，可以分为液相色谱 (LC) 及气相色谱法 (GC) 两大类。

固定相可以是固体，也可以是液体，但是这个液体必须附载在某个固体物质上，这个固体物质称载体或担体（support）。按固定相所处状态不同，液相色谱法又可分为液固色谱法（LSC）及液液色谱法（LLC）。气相色谱法又可分为气固色谱法（GSC）及气液色谱法（GLC）。

（二）按色谱的分离机制分类

按照色谱作用分离机制的分类方法是最重要和最基本的分类方法，通常可分为下面几种。

1. 吸附色谱法（adsorption chromatography）

固定相是固体吸附剂，利用各组分在吸附剂表面吸附能力的差别而分离。当混合物随流动相通过由吸附剂组成的固定相时，由于吸附剂对不同组分有不同的吸附力，从而不同组分随流动相移动的速度不同，最终可将混合物中不同组分分离。这种分离方法取决于待分离物质被吸附剂固定相所吸附的能力，以及它们在分离时所用的溶剂流动相中的溶解度这两个方面的差异。

2. 分配色谱法（partition chromatography）

固定相是一种被担体固定住的液体，利用各组分在固定相与流动相中的溶解度不同，引起分配系数的差别而分离。LLC 与 GLC 都属于分配色谱法。流动相的极性大于固定相极性的液相色谱法，称为反相（reversed phase，RP）色谱法，反之，称为正相（normal phase，NP）色谱法。

3. 离子交换色谱法（ion exchange chromatography，IEC）

固定相是离子交换剂，利用各组分对离子交换剂的可交换基团交换能力（交换系数）的差别而进行分离的方法。

4. 凝胶色谱法（gel chromatography）

固定相是一种多孔性凝胶，利用各组分的分子大小不同，因而在凝胶上受阻滞的程度不同而得到分离。凝胶色谱法也称分子筛色谱法，凝胶起到分子筛的的作用，对分子大小不同的组分起过滤作用，故凝胶色谱又称为凝胶过滤。

5. 亲和色谱法（affinity chromatography，AC）

将具有生物学活性（如酶、辅酶、抗体等）的配位基键合到载体或基质表面形成固定相，利用生物大分子与亲和色谱固定相配位基的亲和力进行分离的色谱法，称为亲和色谱法。

6. 化学键合相色谱法（chemical bonded phase chromatography，BPC）

将固定相的功能基团键合在载体表面，形成的固定相称为化学键合相。使用这种固定相的色谱法称为化学键合相色谱法，简称键合相色谱法。化学键合相可作为液-液分配色谱、离子交换色谱、亲和色谱、高效液相色谱等色谱法的固定相。

7. 毛细管电泳法（capillary electrophoresis，CE）

又称高效毛细管电泳（HPCE），是一类以毛细管为分离通道、以高压直流电场为驱动力的新型液相分离分析技术。它实际上包含色谱、电泳及其交叉内容，其分离机制是靠色谱与电场两种作用，依据样品组分的分配系数及电泳迁移率的差别而分离。毛细管电泳法在 20 世纪 80 年代后期迅速发展起来，它的理论塔板数可高达 10^6/m、灵敏度极高、应用范围广泛，是高效液相色谱之后的又一重大进展。由于毛细管电泳通常采用色谱中的概念或名词，所以一般都放在色谱技术中进行讨论。

三、常用的色谱方法

（一）凝胶色谱

1. 原理

凝胶色谱（gel chromatography）是指混合物随流动相流经装有凝胶作固定相的色谱柱时，混合物中各种物质因分子大小不同而被分离的技术。凝胶，从广义上讲是指一类具有三维空间多孔网状结构的物质，如琼脂糖凝胶、交联葡聚糖凝胶等。由于凝胶色谱过程与过滤相似，故又名凝胶过滤或分子筛过滤；由于物质在分离过程中的阻滞减速现象，故亦称为阻滞扩散色谱或分子排阻色谱。

凝胶色谱的机理是分子筛效应，如同过筛那样，它可以把物质按分子大小不同进行分离，但这种“过筛”与普通的过筛不一样。凝胶色谱柱中装填的是许多直径小于1mm的凝胶颗粒，凝胶颗粒内部又是由三维多孔网状结构构成。在洗脱过程中，当分子大小不一的混合物样品加到凝胶床面后，大分子物质因其分子直径大于凝胶网孔而不能进入凝胶颗粒内部，只能沿着凝胶颗粒的空隙随溶剂流动，受到的阻滞作用小，流程短而移动速度快，先流出色谱柱；分子量小的组分则由于其分子直径小于网状结构的孔径而进入凝胶颗粒内部，受到网孔的阻滞作用，流程长而移动速度慢，比大分子物质后流出色谱柱，从而使分子大小不同的混合物得到分离。图1-11可以形象地看到这个分离过程的实质。

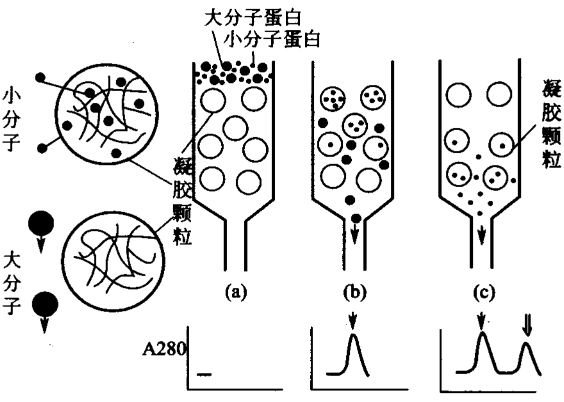


图1-11　凝胶色谱原理及分离过程

（a）蛋白质混合样品加样；

（b）样品上柱后进行洗脱，小分子物质所受阻滞作用大，落在后面，大分子物质先洗脱下来；

（c）小分子物质后洗脱下来

2. 凝胶

凝胶是由胶体溶液凝结而成的固体颗粒状物质，其内部都具有很微细的多孔网状结构。目前市场上供应的色谱用凝胶主要有交联葡聚糖凝胶、交联聚丙烯酰胺凝胶以及琼脂糖凝胶等。

（1）交联葡聚糖凝胶　瑞典出品商品名称为Sephadex，国产商品名称为Dextran，它是由葡聚糖（右旋糖酐）和甘油通过醚桥（$—O—CH_2—CHOH—CH_2—O—$）相交联而成的多孔性网状结构。由于交联度的不同，Sephadex可以分成不同的型号。交联度大的孔隙小，吸水少，膨胀也少，用于小分子量物质的分离；交联度小的孔隙大，吸水多，膨胀也大，适用于大分子物质的分离。交联度常用“吸水量”表示，即每克干凝胶所吸收的水分质量，用这个量比较交联度的大小。商品凝胶的型号采用“吸水量”（水容值）的10倍数字来表示。例如每克凝胶吸水量为2.5g即定为G-25型。

（2）聚丙烯酰胺凝胶　聚丙烯酰胺凝胶是一种人工合成的凝胶，美国出品的商品名称为Bio-gel P，为颗粒状干粉，在溶剂中能自动吸水溶胀成凝胶。聚丙烯酰胺是由丙烯酰胺与亚甲基双丙烯酰胺共聚交联而生成的凝胶，经干燥粉碎或加压成形处理即成Bio-gel P。控制单体和交联剂的比例可得到不同型号的Bio-gel P。Bio-gel P有P-2～P-300等10种，命名是

根据分子量的近似排斥限度。所谓排斥限度是指凝胶能将物质排斥在其颗粒外的分子大小的最低限度，以分子量表示。例如 Bio-gel P-6，其排斥限度分子量为 6×10^3。

（3）琼脂糖凝胶　琼脂糖凝胶是 D-半乳糖和 3,6-（脱水）-2-半乳糖相间结合的链状多糖，瑞典出品的商品名称为 Sepharose，此种产品通常有 2B、4B 两种，分别代表含琼脂糖凝胶 2%、4%的凝胶，凝胶浓度越低其结构越松散，即多孔性越高。美国出品的琼脂糖凝胶的商品名为 Bio-gel A。

（4）Superdex 凝胶　Superdex 凝胶是最新的凝胶填料，它是将葡聚糖以共价方式结合到高交联的多孔琼脂糖珠体上形成的复合凝胶。琼脂糖的高交联骨架增强了珠体的理化稳定性，葡聚糖为介质提供了优良的选择性。这类凝胶物理化学性质稳定、刚度强，适合于高流速，且分辨率高，使用 pH 范围为 3～12；在 pH7 时可经 120℃ 消毒处理，在 1mol/L NaOH、0.1mol/L HCl、1%SDS、6mol/L 盐酸胍和 8mol/L 尿素中使用，色谱行为不变。目前有 Superdex G-30、Superdex G-75 和 Superdex G-200 可供选择，它们的分离范围（相对分子质量）对于蛋白质来讲，分别为 10000、70000 和 600000。

（二）离子交换色谱

离子交换色谱（ion exchange chromatography，IEC）是目前在生物大分子提纯中得到最广泛应用的方法之一。它是以离子交换剂为固定相，利用不同的生物大分子的带电部分与具有相反电荷的离子交换剂吸附强弱不同，而将混合物中的不同离子进行分离的色谱技术。

1. 原理

离子交换剂是一类具有特殊网状立体结构的高分子多元酸或多元碱的聚合物，不溶于水和许多有机溶剂。聚合颗粒中带电荷的酸性或碱性基团作离子交换基团，通过静电作用与带相反电荷的离子（反离子）结合。当流动相中存在带相反电荷的离子时，就与先结合在固定相交换基团上的反离子进行交换。根据可交换离子的性质，离子交换剂分为阳离子交换剂和阴离子交换剂两大类。阳离子交换剂分子中具有酸性基团，能和流动相中的阳离子进行交换。例如：

$$R{-}SO_3^- H^+ + Na^+ \rightleftharpoons R{-}SO_3^- Na^+ + H^+$$

阴离子交换剂分子中具有碱性基团，能和流动相中的阴离子进行交换。例如：

$$R_4{-}N^+ OH^- + Cl^- \rightleftharpoons R_4{-}N^+ Cl^- + OH^-$$

离子交换剂分离样品时，主要依靠增加流动相的离子强度或改变酸碱度来进行洗脱。

（1）增加流动相的离子强度　当缓冲溶液的离子强度增加时，即增加了它与生物大分子对交换基团竞争吸附的能力，把生物大分子置换下来。

（2）改变酸碱度　当缓冲溶液的 pH 接近生物大分子的等电点时，生物大分子的净电荷为零，从而可被洗脱。

由于不同生物大分子的电荷密度分布不同、电荷量不等、等电点的差异以及分子大小的区别，因而与离子交换剂的结合强度不同。可以先把样品交换到柱上，然后逐渐改变置换条件（提高流动相中的离子强度或改变 pH），把各组分从离子交换柱上依次洗脱下来，达到分离纯化的目的。

2. 离子交换剂的类型

常用的离子交换剂为人工合成的有机物，包括离子交换树脂、离子交换纤维素、离子交换葡聚糖和离子交换琼脂糖等。根据交换基团酸碱性的强弱，阳离子交换剂又分为强酸型和弱酸型，阴离子交换剂又为强碱型和弱碱型。在实际工作中，可根据被分离物的性质选用适

当类型的离子交换剂。如羧甲基纤维素（CM-纤维素）常用来分离中性或碱性蛋白质；二乙氨基乙基纤维素（DEAE-纤维素）常用来分离中性或酸性蛋白质；而离子交换树脂常用于分离小分子物质（如氨基酸、核苷、核苷酸等）及制备去离子水。常用离子交换剂的类型见表 1-7。

表 1-7 常用的离子交换剂的种类及解离基团

商品名	解离基团	类别	商品名	解离基团	类别
DEAE	二乙基氨基乙基	弱阴离子	CM	羧甲基	弱阳离子
QAE	氨基乙基	强阴离子	SE	磺酸乙基	强阳离子
TEAE	三乙基氨基乙基	阴离子	SP	磷酸根	强阳离子

（1）离子交换树脂　离子交换树脂是一种特殊网状分子结构的不溶性高分子化合物，大多数离子交换树脂的母体或原料是不溶性的交联聚苯乙烯、交联聚甲基丙烯酸、酚醛、硅酸铝钠等，通过化学反应把带电的基团连接到母体上。带负电的具有结合阳离子的功能，称为阳离子交换树脂；带正电的具有结合阴离子的功能，称之为阴离子交换树脂。离子交换树脂的种类很多，常用的有 Dowex 1 系列阴离子交换树脂、Dowex 50W 系列阳离子交换树脂、国产磺酸型阳离子交换树脂（732 型）等。

（2）纤维素类离子交换剂　纤维素离子交换剂为早期的分离介质之一，而且一直沿用至今，一般分为纤维状、微粒、短纤维及球状四种类型。纤维素离子交换剂表面积较大，开放性的支持骨架使大分子能自由地进入和快速地扩散，离子-纤维素上的交换基团排列松散，对生物大分子的吸附不太牢固，用温和的洗脱条件就可以将其洗脱下来，因此不至于引起生物大分子的变性，且回收率高；多用于蛋白质、核酸、酶、寡核苷酸、肽及病毒等的分离纯化和分析工作。生物大分子的等电点一般在中性附近，因而 DEAE-纤维素和 CM-纤维素应用最广。当要求在较强的酸性条件下（pH＜3）分离生物大分子时，可用 P-纤维素或 SE-纤维素；而在较强碱性条件下（pH＞10）分离时，可用 GE-纤维素。

（3）葡聚糖（Sephadex）类离子交换剂　离子交换葡聚糖为 Pharmacia 公司以 G 系葡聚糖凝胶 G-25 和 G-50 两种凝胶过滤介质，与四种不同的功能基团连接而成的葡聚糖离子交换剂。分为 A 型阴离子交换剂和 C 型阳离子交换剂，分别标以 A-25 或 A-50，C-25 或 C-50。离子交换葡聚糖与离子交换纤维素不同的是交联葡聚糖的骨架为三维网状结构，使得这种离子交换剂不仅具有离子交换的作用，同时还具有分子筛的作用。离子交换葡聚糖由于其单位质量所携带的大量离子交换基团，使得其总交换容量为纤维素类的 3～5 倍。但是，离子交换葡聚糖具有床体积易受 pH 和盐浓度影响而变化的缺点，对分离的结果有一定的影响。

（4）琼脂糖类离子交换剂　这类离子交换填料是在琼脂糖表面进行化学改造而制备的。由于琼脂糖表面含有大量的羟基，化学改造极为方便。目前在生物大分子的分离中绝大多数采用琼脂糖系列的离子交换填料。

离子交换色谱通常采用柱色谱。柱色谱过程包括离子交换剂的选择、离子交换剂使用前的处理与变型、装柱、样品的准备与加样、洗脱、检测及离子交换剂的再生等若干步骤。

经过离子交换被吸附在离子交换剂上的待分离物质，有两种洗脱方法：一是增加离子强度，使洗脱液中的离子能争夺交换剂的吸附部位，从而将待分离的物质置换下来；二是改变 pH，使样品离子的解离度降低，电荷减少，因而对交换剂的亲和力减弱而被洗脱下来，如低 pH 洗脱液容易使阴离子交换剂的样品洗脱下来。

从洗脱液的成分而言，洗脱方式有两种：一是选用几种洗脱能力逐步增强的洗脱液相继洗脱，此为分步（分段）洗脱法，适用于各组分对交换剂亲和力比较悬殊的样品；二是选用离子强度和 pH 呈连续梯度变动的洗脱液进行洗脱，使洗脱能力持续增强，此为连续梯度洗脱法，适用于各组分与交换剂亲和力相近的样品。

（三）亲和色谱

亲和色谱利用生物大分子之间的特异而可逆的相互作用而进行分离与分析，这种方法也叫做生物亲和色谱或生物特异性亲和色谱。这种可以特异可逆结合的物质很多，如抗原与抗体、底物与酶、激素与受体等，它们间的这种特异亲和能力又叫亲和力。

亲和色谱中，一对互相识别的分子互称对方为配体，如激素可认为是受体的配体，受体也可以认为是激素的配体。亲和色谱法是将配体固定在固相载体上，并且将其装入色谱柱中，当样品通过色谱柱时，由于配体和相对应的生物大分子分子有专一性的亲和作用，将某种生物大分子吸附在柱中，样品中的其他组分不产生这种专一性的结合，而直接流出色谱柱。然后，便可以利用洗脱剂将吸附在柱中的生物大分子洗脱下来。亲和色谱法具有高度的专一性，而且色谱过程简单、快速，是一种理想的有效分离纯化生物大分子的手段。

1. 固相载体的选择

对于一个成功的亲和色谱分离来说，一个重要的因素就是选择合适的固体载体。一个理想的载体，首先它必须尽可能少地同被分离的物质进行相互作用，以避免非特异的吸附作用。因此，优先选用的是中性聚合物，例如，琼脂糖（Sepharose）或聚丙烯酰胺凝胶（Bio-gel）。其次，载体必须具有良好的通透性，即使在亲和剂键合在它的表面之后也必须保持这种特性。连接亲和剂的先决条件是有足够量的某些化学基团存在，这些基团在不影响载体的结构，也不影响被连接的亲和剂的条件下被活化或衍生。载体在结合亲和剂后，必须在机械性能和化学性质上具有稳定性，而且在改变 pH、离子强度、温度以及变性剂的条件下也应该稳定。载体必须有大的孔网结构，允许大分子物质自由出入。一般常用的载体有纤维素、葡聚糖凝胶、琼脂糖凝胶、聚丙烯酰胺凝胶和多孔玻璃等。

2. 配体的选择

正确地选择合适的配体以及合适的结合方式，对获得具有优良分离效果和较大容量的载体同样具有重要作用。能与分离物质牢固、特异和可逆结合的物质都可以作为配体。选择配体有两个条件。第一是生物大分子与配体间具有合适的亲和力，亲和力太强，洗脱条件剧烈，易造成生物大分子失活；亲和力太小，解离容易，结合率不高。第二是配体要具有双重功能，既有可牢固与载体结合的基团，结合后又不影响生物大分子与配体间的亲和力。亲和色谱要取得成功，配体结合在载体上并不是唯一的因素。具有同等重要的是要完全地从载体上除去非共价结合的配体。因此，应当十分认真地清洗并检查清洗的情况。

3. 配体与载体之间的偶联

实践证明，琼脂糖凝胶和聚丙烯酰胺凝胶是亲和色谱的优良载体。琼脂糖凝胶结构开放，通过性好，酸碱处理时相当稳定，物理性能也好。琼脂糖凝胶上的羟基在碱性条件下极易被溴化氰活化成亚氨基碳酸盐，并能在温和的条件下与氨基等基团作用而引入配体，亲和色谱中最为常用的琼脂糖凝胶的型号是 Sepharose-4B。琼脂糖被溴化氰活化后，再与生物大分子结合形成亲和色谱固定相是亲和色谱中最为常用的一种。

4. 使用方法

亲和色谱的分离方法随分离的物质不同而异，一般的程序如下。

(1) 吸附　亲和色谱固定相制备好后，装入色谱柱中，色谱柱无特殊要求，常用短而粗的柱子。样品要通过透析等方法将溶液转换为起始缓冲溶液。上样量可根据柱子的吸附容量来推算，通常为吸附容量的1/3或更低。在样品上柱后，使用10倍柱体积的缓冲溶液将不结合的杂质清洗掉。

(2) 洗脱　从柱中洗脱目标产物是亲和色谱是否成功的关键，通常采用降低目标产物与配体之间的亲和力的方式进行洗脱。可用改变流动相离子强度（浓度）的方式将目标产品洗脱下来，也可以采用改变流动相pH的方法将目标产品洗脱下来。对于有些吸附十分牢固的目标产物，有时只有采用较强的酸或碱，甚至变性剂来作为洗脱剂，这有使生物样品失活的危险，需立即对洗脱的蛋白质溶液进行中和、透析或稀释。

(四) 高效液相色谱

高效液相色谱（high performance liquid chromatography，HPLC）又叫高压（高速）液相色谱。它在生物大分子的分离和纯化方面占据了极其重要的地位。在近几年发表的色谱文章中，约有80%以上的生物大分子是由HPLC进行纯化的。表1-8列出了经典柱色谱与HPLC在分离生物大分子方面的差别。

表1-8　经典柱色谱与HPLC在分离生物大分子时差别

性　能	经典柱色谱	HPLC
填料粒度	>100μm	2～10μm(小的可到1.5μm)
柱效	低	高，(4～6)万塔板数/m
分离组分数	几个	几十
柱子大小	(1～2.5)cm×(10～100)cm	(0.2～1.0)cm×(5～25)cm
压力	几百千帕	40～50兆帕(MPa)
分离时间	几小时到几十小时	几分钟到几十分钟
自动化程度	人工操作	微机控制，在线检测
柱子寿命	使用几次或数十次	连续使用数百次以上

HPLC在分离生物大分子时同经典色谱一样，按照分离机制，可将其分为：高效反相色谱（HFRPC）、高效疏水作用色谱（HPHIC）、高效离子交换色谱（HPIEC）、高效排阻色谱（HPSEC）和高效亲和色谱（HPAC）等。其中高效反相色谱的应用为最多。它们的主要特性见表1-9。

表1-9　生物大分子在不同HPLC上的分离机制及其操作条件间的关系

色谱类型	分离机制	操作条件	
		固定相上的基团	流动相
高效反相色谱(HPRPC)	疏水作用	非极性	有机溶剂水
高效疏水作用色谱(HPHIC)	疏水作用	极性、非极性相间	盐水溶液
高效离子交换色谱(HPIEC)	静电作用和电荷分布	带电荷的阴、阳离子	盐水溶液
高效排阻色谱(HPSEC)	分子大小和形状	惰性	盐水或有机溶剂水
高效亲和色谱(HPAC)	分子间特异的亲和力	亲和配体	盐水溶液

1. 高效反相色谱（HPRPC）

反相色谱法是一个通用性的蛋白质分离技术。它是利用蛋白质表面的非极性区域同固定相表面的非极性配体相互作用，按照疏水作用力大小进行分离的一种方法。在分离时，HPRPC由于使用酸性极性有机溶剂作为流动相，使用疏水性极强的非极性烷链作为固定相的配体，使得蛋白质在HPRPC上进行分离时常常失去生物学活性，所以使用HPRPC进行

蛋白质分离时常常受到限制。

2. 高效疏水作用色谱（HPHIC）

HPHIC 因其廉价和纯化后保持高的蛋白质生物学活性而成为一个较新的蛋白质分离技术。它的分离是依据蛋白质与固定相间的疏水作用不同而进行的。洗脱体系为高浓度的盐水溶液。当在高盐浓度时，蛋白质被吸附在固定相的表面，随着盐浓度的降低，蛋白质逐渐地被分离开来。

HPHIC 使用的流动相为高浓度的盐水溶液（最常用的盐水体系为硫酸铵-水溶液），对蛋白质的构象具有一定的稳定作用，所以在分离后，蛋白质可保持高的活性回收率。目前，利用 HPHIC 分离的蛋白质有糖蛋白、单抗、酶和基因重组蛋白质等。

3. 高效离子交换色谱（HPIEC）

HPIEC 是蛋白质纯化中最为常用的色谱法，占各种蛋白质色谱分离技术使用量的 75%左右。它是依据蛋白质与分离填料间的静电作用力不同而进行分离的一种方法。HPIEC 在分离蛋白质时，由于采用的流动相为低浓度的盐水体系，使得蛋白质在分离后的活性回收率大于 90%。

4. 高效排阻色谱（HPSEC）

HPSEC 是蛋白质分离的另外一种方式，它是按照蛋白质的大小序列进行分离的一种技术。不涉及太多的作用力问题，操作简单。作为蛋白质分离的一种方法，它具有如下优点。

① 分离条件温和，蛋白质的活性回收率高，大部分蛋白质在 HPSEC 分离后的活性回收率大于 90%。

② 蛋白质的分离是按分子的大小顺序进行分离的，不需要使用梯度分离方法，操作简单，对于未知的蛋白质在分离后就可以估算出它的相对分子质量的大小。

③ HPSEC 是蛋白质除盐最为常用的方法，比透析方法快。

但 HPSEC 最大的缺点在于分离容量小，蛋白质在分离时的上样体积通常为柱体积的 1%左右，最大不能超过柱体积的 5%，因此在分离时通常将它放在纯化工艺的最后。由于 HPSEC 在分离过程中蛋白质的保留时间不会超过柱子的外水空间，所以，柱效较以上的色谱方法都低。

5. 高效亲和色谱（HPAC）

HPAC 分离技术目前已成为高效色谱分离领域中的一个重要分支，它在分离过程中表现出的高选择性和特异性是其他任何技术无法比拟的，但 HPAC 技术在发展中也遇到某些固有的限制，如不易在大规模生产中使用、无法直接从浑浊溶液中分离目标产物等不足。为了克服亲和色谱技术本身的缺陷，近几年人们对其进行了大量的改造，使其得到了巨大的发展，新的亲和分离技术也不断产生，如磁性亲和分离技术、径向亲和分离技术、膜亲和分离技术、双水相亲和分离技术、亲和沉淀分离技术、融合蛋白标签技术等。

（五）毛细管电泳

毛细管电泳（capillary electrophoresis，CE）又称高效毛细管电泳（HPCE），是一类以毛细管为分离通道，以高压直流电场为驱动力的新型液相分离、分析技术。它实际上包含色谱、电泳及其交叉内容，其分离机制是靠色谱与电场两种作用，依据样品组分的分配系数及电泳的迁移率差别而分离。毛细管电泳法是 20 世纪 80 年代中后期迅速发展起来的，它的柱效可高达 10^6/m，灵敏度极高，应用范围广泛，在分析技术中，是高效液相色谱之后的又一重大进展。由于毛细管电泳通常采用色谱中的概念或名词，所以一般都放在色谱技术中进行讨论。

1. 毛细管电泳仪

毛细管电泳仪是由毛细管、高压电源、电极及电极槽、毛细管进样装置、填灌清洗装置、恒温装置、在线检测器、数据收集和处理系统等几部分组成，其基本结构如图 1-12 所示。

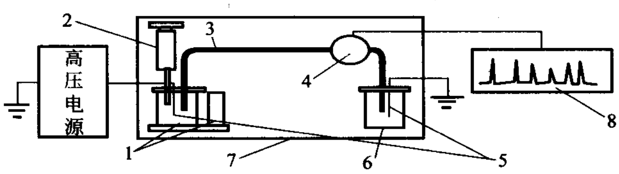


图 1-12　毛细管电泳系统的基本结构图

1—高压电极槽与进样装置；2—填灌清洗装置；3—毛细管；4—检测器；
5—电极；6—低压电极槽；7—恒温装置；8—数据处理/记录显示装置

(1) 高压电源和进样装置　毛细管电泳仪的高压装置一般采用 0～30kV 连续可调直流高压电源，最大电流为 200～300μA，并且可改变其正、负极方向。进样方式主要分为电动法、压力法和扩散法。

(2) 毛细管填灌清洗装置　采用正压或负压设备可实现对毛细管的清洗和填灌。

(3) 毛细管　毛细管一般是由熔融石英制成，内径通常为 25～75μm，其外壁涂有一层聚亚胺（polyimide）保护层，使得毛细管不易折断，长度为 50～70cm，容积仅几微升。

(4) 检测器　要求检测器具有高灵敏度和低噪声，折射指数、电化学、紫外/可见光吸收、荧光、激光诱导荧光、质谱等检测方法均被应用于毛细管电泳。其中紫外/可见光吸收最为常用，检测限为 (10^{-5}～10^{-6})mol/L；激光诱导荧光的灵敏度较高，检测限为 (10^{-10}～10^{-12})mol/L，但价格昂贵。

(5) 电极　正、负电极均为铂金丝。

(6) 低压电极槽　电极槽通常是带螺口的玻璃瓶，便于密封。

(7) 恒温装置　高压电泳过程会产生大量的热，导致分离效率和重现性变差，常用液冷或气冷设备把温度控制在±0.1℃。

(8) 数据处理/记录显示装置　与色谱要求相同，一般采用计算机进行控制处理。

2. 毛细管电泳中的电渗

当固体与溶液接触时，大多数固体表面（例如石英毛细管内表面）都带有多余（额外）的负电荷，溶液中的正离子吸附至固体（石英）表面形成双电子层，当在毛细管两端施加了电压后，这一层正离子就向负极移动，并带动毛细管中的溶液以液流的形式移向负极，这一现象称为电渗。

由于毛细管的表面积与体积之比大，且电泳时使用了高电压，电渗流在毛细管电泳中常常起着不可忽视的作用。电渗流是伴随电泳而产生的一种电动现象，其速度约为一般离子电泳速度的 5～7 倍。所以，电渗流在毛细管电泳中起到 HPLC 中恒流泵一样的作用，在一次电泳操作中可以同时完成阳离子、阴离子的分离，不带电荷的中性分子也能随电渗流一起移动。

电渗流的方向决定于毛细管内壁表面电荷的性质，一般情况下，石英毛细管内壁表面带负电荷，电渗流的方向为由阳极到阴极。电渗流的特点之一是液体沿着毛细管壁均匀流动，因而其前沿是平齐的，不同于高效液相色谱中泵推动的液体流动呈圆弧状（见图 1-13）。这也使得毛细管电泳比 HPLC 有更高的分辨率。

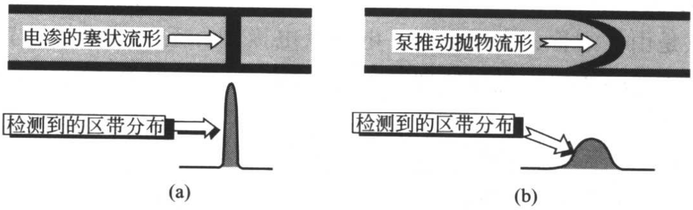


图 1-13　毛细管电泳的电渗流与 HPLC 流动相的流型以及它们引起样品的区带展宽
(a) 毛细管电泳的电渗流驱动流型；(b) HPLC 的压力驱动流型

3. 进样

毛细管电泳有 3 种进样方式。第一种为压力进样，即在毛细管进样端施加一定压力推入样品，或在出口端连接真空装置吸取样品。其优点是进样体积和样品浓度、样品溶液离子强度等基本无影响。第二种为电动进样，即在一定的电场下，样品通过电渗流及电迁移进入毛细管中。这种方式的进样量受电渗、样品浓度、样品分子的泳动力以及样品溶液中无机离子的多少及种类等多种因素影响，因此其进样重复性不如前者，但是该方法可实现完全自动化操作，而且当毛细管中填充了黏性较大的介质或凝胶时，只能选择电动进样方式。第三种为扩散进样，即将毛细管插入样品溶液时，样品分子因浓度差而向毛细管内扩散。扩散进样动力属不可控制参数，进样量仅由扩散时间来控制。毛细管电泳的进样体积可小到 1μl 以下，消耗样品体积仅 1～50nl。

4. 电泳和检测

由于毛细管具有很好的散热性，允许在很高的电场强度下（可达 400V/cm 以上）进行电泳，因此可在很短的时间内（几分钟）完成高效分离。毛细管电泳通常采用在线检测的方式，在石英毛细管长度的 3/4 处利用紫外/可见光或激光透过毛细管来检测分离的结果。电泳缓冲液必须在紫外光下吸收低，否则会增加基线噪声并降低检测信号。与检测器相连的记录系统，除可显示分离图谱外，还可以根据峰面积的积分进行定量分析。

5. 毛细管电泳分离的几种模式

毛细管电泳分离的模式很多，比较常用的有下面几种。

(1) 毛细管区带电泳　毛细管区带电泳（capillary zone electrophoresis，CZE）是指溶质在毛细管内的背景电解质溶液中以不同速度迁移而形成一个一个独立的溶质带的电泳模式，其分离基础是迁移率的差别。毛细管区带电泳的突出特点是简单，CZE 模式是 HPCE 中最简单、应用最广的一种操作模式，是其他操作模式的基础。

(2) 毛细管凝胶电泳　毛细管凝胶电泳（gel capillary electrophoresis，GCE）是毛细管内填充凝胶或其他筛分介质，如交联或非交联的聚丙烯酰胺。电荷/质量之比相等但分子的大小不同的分子，经电场力的推动，在凝胶聚合物构成的分子筛中电泳，其运动受到网状结构的阻碍。大小不同的分子经过分子筛时受阻力不同，大分子受到的阻力大，在毛细管中迁移的速度慢；小分子受到的阻力小，在毛细管中迁移的速度快，从而使它们得以分离。

(3) 毛细管等电聚焦　毛细管等电聚焦（capillary isoelectric focusing，CIEF）是一种根据等电点差别分离生物大分子的电泳技术。两性物质以电中性状态存在时的 pH 叫等电点，用 p*I* 表示。氨基酸、蛋白质、多肽等在其等电点时为电中性，迁移率为零，溶解度最小。当两性电介质在毛细管中形成 pH 梯度时，不同等电点的溶质在外电场的作用下，分别向其等电点的 pH 范围迁移，此过程叫聚焦。当溶质迁移到等于其等电点的 pH 范围时，就

不再移动，形成一个很窄的溶质带，不同的两性物质由于等电点不同，因此相互分离开来。

（4）毛细管等速电泳　毛细管等速电泳（capillary isotachophresis，CITP）用两种迁移率差别大的缓冲体系分别构成前导离子和尾随离子，将样品组分像夹心饼干一样夹在二者之间得到分离。一般前导电解质用迁移率大于试样中所有负离子的电解质组成，尾随离子由迁移率小于试样中所有负离子的电解质组成。所有溶质都按前导离子的速度等速前移。由阴极进样，阳极检测。电泳时，所有负离子都向阳极迁移。前导离子迁移率最大，迁移最快，走在最前，其后是迁移率次之的负离子，它们都以前导离子速度迁移，并逐渐形成独立的溶质区带而得到分离。

（5）毛细管电色谱　毛细管电色谱（capillary electro chromatography，CEC）是把毛细管电泳和毛细管液相色谱结合起来的一种分离技术，采用装有固定相的毛细管柱（或有固定相涂层的空心毛细管柱），以电渗流为驱动力，使样品在两相间进行分配的色谱。普遍使用的是ODS的毛细管填充柱（50～100μm），让样品在固定相和流动相之间反复进行分配而得到分离。

6. 毛细管电泳与其他分析方法的比较

毛细管电泳是基于电泳的高分辨机制和色谱仪器自动化概念而诞生的，因而兼具二者的优越性，即分辨率高、分析速度快、消耗样品少、灵敏度高、操作简便、分离模式多样、不使用毒性强的有机溶剂等诸多优点。表 1-10 列出毛细管电泳与其他分析方法的比较。

表 1-10　毛细管电泳与其他分析方法的比较

比较项目	毛细管电泳	凝胶电泳	HPLC
分离机制	电荷、疏水性、分子大小、p*I*	电荷、分子大小、p*I*	极性、分配系数、分子大小、电荷、亲和力
常规分析时间	10～30min	2h～几天	10～120min
分辨率	优秀	优秀	良好
样品需要量	<1μl	1～100μl	10～50μl
灵敏度(常规分析)	pg～ng	ng～μg	ng～μg
检测方式	在线检测，可见/紫外、电化学、激光诱导荧光等	非在线检测，染色、放射性显影	在线检测，可见/紫外、荧光、电化学等
仪器价格	较高	较低	较高
实验消耗	很低	较高	较高
自动化程度	高	低	高
制备能力	无	有	高
部分收集	困难	困难	容易

目前毛细管电泳技术发展迅速，具有很好的应用前景。但是它只能局限于分析水平，尚不能作为分离制备的手段，而且它还不能分析难溶于水的样品，因此不能完全替代气相色谱、液相色谱及传统的凝胶电泳。

第四节　离　心　技　术

离心技术是利用离心机高速旋转时产生强大的离心力，使置于离心机转头悬浮液中的颗粒发生沉降或漂浮，从而使这些颗粒浓缩或与其他颗粒分离的方法。离心机的种类繁多，用

途各异，本节只介绍生物离心机的基本原理、方法和应用。

一、离心理论

（一）离心分离的原理

需要离心分离的生物样品常预先制成悬浮液。将处于悬浮状态的细胞、细胞器、病毒和生物大分子等都称为“颗粒”。每个颗粒都有一定大小、形状、密度和质量。当离心机转子高速旋转时，这些颗粒在介质中发生沉降或漂浮，它们的沉降速度与作用在颗粒上的力的大小和方向有关。颗粒除受到离心力（F_c）外，还受到颗粒在介质中移动时的摩擦阻力（F_f）、与离心力方向相反的浮力（F_B）、颗粒处于重力场之下的重力（F_g）和与重力方向相反的浮力（F_b）。各力的作用方向见图 1-14。此外，还有周围介质小分子的作用力，当颗粒很小时，分子对颗粒的作用力十分明显，要使这种小介质颗粒沉降，需要更大的离心力。本节只讨论比介质分子大得多的颗粒，因此介质作用力不予考虑。下面将对各个力作详细的分析。

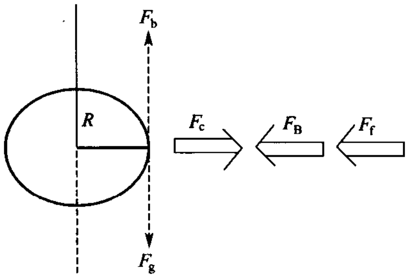


图 1-14　颗粒离心时所受作用力示意图
F_c：离心力；F_B：浮力；F_f：摩擦阻力；F_g：重力；F_b：由重力引起的浮力

1. 离心力

离心力（F_c）的大小等于离心加速度 ω^2R 与颗粒质量 m 的乘积，即：

$$F_c = m\omega^2 R \tag{1-1}$$

式中，ω 是旋转角速度（rad/s）；R 是颗粒离开旋转中心的距离（cm）；m 是质量（g）。

这里必须指出，根据离心机应用的习惯，将颗粒离开轴心的方向定为离心力的方向。实质上这与物理学上所指的向心力方向正好相反。

2. 重力及重力浮力

重力（F_g）是颗粒质量与重力加速度的乘积，用下式表示：

$$F_g = m \times g \tag{1-2}$$

重力的方向与离心力的方向互相垂直，其大小远远小于离心力，可以忽略不计。同时由重力而产生的作用在颗粒上的浮力（F_b）也可忽略不计。

3. 介质的摩擦阻力

Stocke 建议介质对颗粒的摩擦阻力（F_f）用下式表示：

$$F_f = 6\pi\eta \cdot r_p \frac{dR}{dt} \tag{1-3}$$

式中，η 是介质的黏滞系数（厘泊，cP）；r_p 是颗粒的半径（cm）；dR/dt 是颗粒在介质中的移动速度（cm/s），又称为沉降速度（V），指单位时间内颗粒沉降的距离。

公式(1-3) 只适用于球形颗粒，但不少生物颗粒并非球形，有椭球形、扁球形、棒形或线形等，这使情况更复杂。对于椭球形颗粒，Stocke 阻力方程应改成为：

$$F_f = 6\pi\eta \cdot r_p \left(\frac{dR}{dt}\right)\frac{f}{f_0} \tag{1-4}$$

式中，f_0 为球形颗粒的摩擦系数；f 为同球形等体积的扁球形或椭球形颗粒的摩擦

系数。

颗粒偏离球形越甚，则阻力 F_f 值也越大。

4. 浮力

在重力场中，浮力的定义是指被物体所排开的周围介质的质量。但在离心场中，颗粒的浮力与离心力方向相反，为颗粒排开介质的质量与离心加速度之乘积。用下式表示：

$$F_B=\frac{m}{\rho_p}(\rho_m)\omega^2 R=\frac{\rho_m}{\rho_p}m\omega^2 R \tag{1-5}$$

式中，ρ_p 为颗粒密度（g/cm^3）；ρ_m 为介质密度（g/cm^3）；m/ρ_p 为介质的体积；$(m/\rho_p)\rho_m$ 为颗粒排开介质的质量。

综上所述，在离心场中，作用于颗粒上的力主要有离心力 F_c、浮力 F_B 和摩擦阻力 F_f。

5. 沉降速度

当离心转子从静止状态加速旋转时，原来处于悬浮状态的颗粒如果密度大于周围介质的密度，则颗粒离开轴心方向移动，即发生沉降；如果颗粒密度低于周围介质的密度，则颗粒朝向轴心方向移动，即发生漂浮。无论沉降或漂浮，离心力的方向与摩擦阻力和浮力方向相反；当离心力增大时，反向的两个力也增大，到最后离心力与摩擦阻力和浮力平衡，颗粒的沉降（或漂浮）速度达到某一极限速度，这时颗粒运动的加速度等于零，速度 dx/dt 变成恒定。那么

$$F_c=F_B+F_f \tag{1-6}$$

将式(1-1)、式(1-3)、式(1-5) 代入式(1-6)，整理后得

$$V=\frac{dR}{dt}=\frac{d^2(\rho_p-\rho_m)}{18\eta\cdot f/f_0}\omega^2 R \tag{1-7}$$

从式(1-7) 可见：①颗粒的沉降速度与颗粒直径的平方、颗粒密度和介质密度的差值以及离心加速度成正比，而与介质的黏滞系数、颗粒偏离球形的程度成反比；②当颗粒的密度 ρ_p 大于介质密度 ρ_m 即 $\rho_p>\rho_m$ 时，颗粒发生沉降；反之，$\rho_p<\rho_m$，颗粒漂浮；而当 $\rho_p=\rho_m$ 时，颗粒不沉不浮；③在离心加速度 $\omega^2 R$ 不变的情况下，颗粒的沉降速度主要决定于颗粒的直径大小和颗粒的形状，而颗粒的密度所起的作用较小。

6. 沉降系数

Svedberg 将沉降系数下的定义为颗粒在单位离心力场作用下的沉降速度，可以下式表示：

$$s=\frac{dX/dt}{\omega^2/R} \tag{1-8}$$

从上式可见，沉降系数可以看作是颗粒沉降速度和离心加速度的比值，沉降系数的单位以 1×10^{-18}s=1 Svedberg 单位，以 1S 表示。例如蛋白质的沉降系数为 4.5S 即 4.5×10^{-18} s，IgG 为 7S 即 7×10^{-18}s 等。沉降系数的物理意义是颗粒在离心力作用下从静止状态到达极限速度所经过的时间。如将式(1-7) 二边同除以 $\omega^2 R$，则得到沉降系数的表示式，在给定的介质中它的大小主要是由颗粒直径 d 的平方和摩擦系数 f/f_0 所决定，见下式：

$$s=\frac{dR/dt}{\omega^2 R}=\frac{d^2(\rho_p-\rho_m)}{18\eta\cdot f/f_0} \tag{1-9}$$

（二）相对离心力和离心时间

1. 相对离心力（RCF）

是指在离心力场中，作用于颗粒的离心力相当于地球引力的倍数，单位是重力加速度

($g=980.6\text{cm/s}^2$)。故 RCF 的公式如下：

$$\text{RCF}=\frac{F_c}{F_g}=\frac{m\omega^2R}{mg}=\frac{\omega^2R}{g}=\frac{(2\pi\cdot\text{r/min})^2R}{3600\times980.6}=1.118\times10^{-5}R(\text{r/min})^2$$

如以 N 代表 r/min，上式为：$\text{RCF}=1.118\times10^{-5}RN^2$ (1-10)

式中，R 为离子转子的半径距离，以 cm 为单位。

一般情况下，低速离心时常以 r/min 来表示，超速离心则以 g 表示。计算颗粒的相对离心力时，应注意离心管与旋转中心的距离 R。由于转子的形状及设计差异，离心管的口部和底部到旋转轴中心距离差异很大。

如：管口 $R=4.8\text{cm}$，管底 $R=8.0\text{cm}$，12000r/min

$\text{RCF}_{口部}=1.118\times10^{-5}\times(12000)^2\times4.8=7734\times g$

$\text{RCF}_{底部}=1.118\times10^{-5}\times(12000)^2\times8.0=12891\times g$

如上所见，作用于离心管口部和底部的离心力相差近乎一倍。这不仅说明了超速离心时用 g 代替 r/min 的原因，也说明了 R 应指旋转轴中心到某被分离物质颗粒在离心管中所处位置的距离，该颗粒所受到的离心力随其在管中的移动而变化。科技文献中，离心力的数据常指其平均值，即在离心管中点处的离心力。

为了便于进行转速和相对离心力之间的换算，在式(1-10) 的基础上制作了 R 与转速、相对离心力三者关系的列线计算图（图 1-15）。由图中两者数值点的连线与第三者的交点，即为所求第三者的数值。

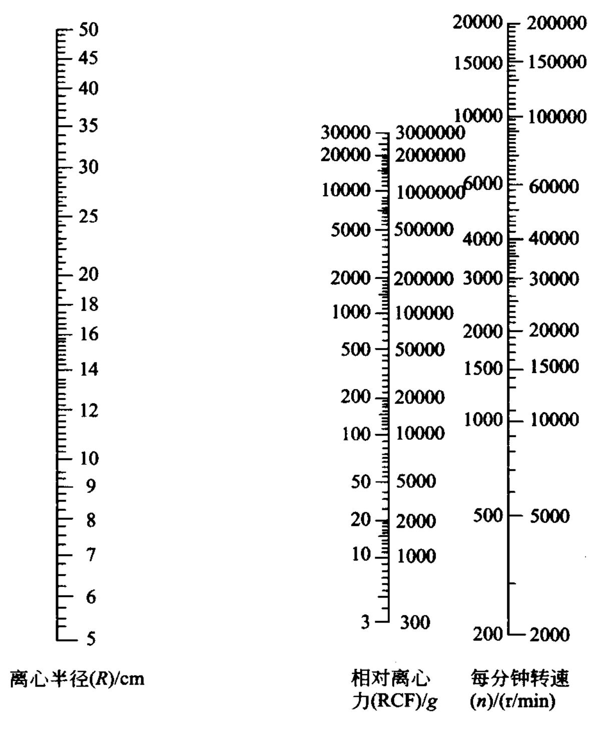


图 1-15 半径与转速、相对离心力关系的列线图

2. 沉降时间（t）

在某一介质中使一种球形颗粒从液体的弯月面沉降到离心管底部所需要的离心时间称为沉降时间。沉降时间与沉降速度成反比。

$$t=\frac{18\eta}{\omega^2 d^2(\rho_p-\rho_m)}\ln\frac{R_{max}}{R_{min}} \tag{1-11}$$

式中，d 为颗粒平均直径（cm），R_{max} 和 R_{min} 代表转轴中心至离心管底部和液面的距离。

如果已知某种颗粒的沉降系数（s），则可估计其沉降时间（t）

$$t=\frac{1}{s}\left[\frac{\ln R_{max}-\ln R_{min}}{\omega^2}\right] \tag{1-12}$$

离心时间是由实验要求所决定，为了避免不稳定颗粒的凝聚、挤压损伤或变性失活，在保证分离的前提下，应尽可能缩短离心时间。相反，分离某些沉降较快的大颗粒时，为了达到预期的分离效果，往往使用黏度较大的介质，以阻止颗粒的过度沉降，并延长离心时间。

（三）离心机的分类

目前在生物医学领域内常用的离心机种类繁多。

按其离心转子能达到的最高转速分类，有：低速离心机（在 6000r/min 以下）、高速离心机（在 25000r/min 以下）、超速离心机（在 30000r/min 以上），它们的性能和应用见表 1-11。目前，商品中的超速离心机最高转速在 100000r/min 左右（达 1000000g）。

表 1-11　三种不同级别的制备离心机的比较

比较项目	普通离心机	高速离心机	超速离心机
最大转速/(r/min)	6000	25000	30000 以上
最大 RCF/g	6000	89000	可达 1000000 以上
分离形式	差速离心法	差速离心法	密度梯度区带离心法或差速离心法
离心管平衡允许误差	0.25g	0.1g	0.1g
转子	角式和外摆式转子	角式、外摆式转子等	角式、外摆式、区带转子等
仪器结构性能和特点	速度不能严格控制，多数室温下操作	有消除空气和转子间摩擦热的制冷装置，速度和温度控制较准确、严格	备有消除转子与空气摩擦热的真空和冷却系统，有精确的温度和速度控制、监测系统，有保证转子正常运转的传动和制动装置
应用	收集易沉降的大颗粒（如红细胞、酵母细胞等）	收集微生物、细胞碎片、大细胞、硫酸铵沉淀物和免疫沉淀物等。但不能有效沉淀病毒、小细胞器（如核糖体）、蛋白质等生物大分子	主要分离细胞器、病毒、核酸、蛋白质、多糖等，甚至能分离分子大小相近的同位素标记物 ^{15}N-DNA 和未标记的 DNA

根据用途不同，超速离心机又可分为制备型超速离心机和分析型超速离心机两大类。制备型超速离心机用于分离和制备较大量的样品，分析型超速离心机附设有特殊的光学系统，可以在离心分离过程中通过光学系统将大分子的沉降行为以扫描或照相形式记录下来，并对其进行分析。近年，出现了制备、分析两用超速离心机，它除了可以安装上普通的制备型转头完成样品的制备工作，还可以通过更换特殊的转子并装上相应的光学分析附件完成对样品的分析工作。

二、离心分离的种类

根据离心原理，可设计出各种离心方法，归纳起来有两大类（表 1-12）。

表 1-12 离心分离法的分类

比较项目	差速离心法	密度梯度区带离心法	
		速度区带离心法	等密度离心法
颗粒沉降系数	相差大(1～n个数量级)	分子量相差3倍以下至相差较少(2%或更少)的蛋白质	—
离心特点	短时间、多次采用不同速度和离心时间进行分段离心	①沉降速度主要依赖于颗粒的形状和大小;②离心时间较短,颗粒沉降速度不可能为零,若时间过长,分离开的区带有可能在管底重新靠拢;③颗粒密度不等于周围介质密度	①沉降平衡主要依赖于颗粒的密度;②离心时间较长,一般大于15h,平衡时沉降速度为零,形成稳定的区带;③沉降平衡时,颗粒密度一定等于周围介质密度
适用范围	分子量相差大,易受梯度介质损伤的颗粒。常用于组织匀浆中分离细胞器及分离病毒	大小不同而密度相似、受差速离心挤压变形、受低浓度梯度介质损伤较少的颗粒。一次分离量较少,宜作分析分离。可分离核酸、蛋白质、核糖体亚基及其他成分(如整细胞、脂蛋白)	大小相同而密度不同,稳定、不受高浓度介质损伤的颗粒。一次分离样品量大,用于制备及分析分离。可分离核酸、亚细胞器、整细胞,也可分离复合蛋白质

（一）差速离心法

差速离心法是通过不断增加相对离心力，使沉降速度不同的颗粒在不同离心速度及不同离心时间下进行多次离心的方法。差速离心法一般用于分离沉降系数相差较大的颗粒。进行差速离心时，首先要选择好颗粒沉降所需要的离心力和离心时间。离心力过大或离心时间过长，容易导致大部分或全部颗粒沉降及颗粒被挤压损伤。先选用较小的离心力和较短的离心时间进行离心，在离心管底部得到最大和最重颗粒的沉淀，分出的上清液进一步加大转速再次进行离心，又得到第二部分较大、较重颗粒的“沉淀”及含更小而轻颗粒的“上清液”。如此反复进行，即能把液体中的不同颗粒较好地分离开。此法所得沉淀是不均一的，仍杂有其他成分，需经再悬浮和再离心（2～3次），才能得到较纯的颗粒。差速离心法主要用于分离细胞器和病毒。

（二）密度梯度离心法

密度梯度离心法包括速度区带离心法和等密度离心法两种方法。后者又可分为预制梯度等密度离心法及自形成梯度等密度离心法两种方法。

1. 速度区带离心法

速度区带离心法是在离心前离心管内先装入密度梯度介质（如蔗糖、甘油、KBr、CsCl等），待分离的样品铺在梯度液的顶部，同梯度液一起离心。由于离心力的作用，颗粒离开原样品层，按不同沉降速度向管底沉降，离心一定时间后，沉降的颗粒逐渐分开，最后形成一系列界面清楚的不连续区带。沉降系数越大，往下沉降越快，所形成的区带也越低；沉降系数较小的颗粒，则在较上部分依次出现。在离心过程中，区带的位置和形状（或宽度）随时间而改变，因此区带的宽度不仅取决于样品组分的数量、梯度的斜率、颗粒的扩散作用和均一性，也与离心时间有关，时间越长，区带越宽。适当增加离心力可缩短离心时间，并减少扩散所导致的区带加宽现象，增加区带面的稳定性。

梯度液的类型有两种，一种是不连续梯度或称阶梯梯度，另一种为连续梯度又叫线性梯度，后者又分直线性梯度及曲线性梯度。下面介绍几种常用的密度梯度材料。

（1）蔗糖　相对分子质量342.3，水溶性大，性质稳定，有较大的黏滞度，渗透压较高，其最高密度可达到1.33g/ml。常用于细胞器、病毒、RNA的分离。由于有较大的渗透压，不宜用于细胞分离。如与甘油制成混合梯度，可用于染色体的分离。

（2）聚蔗糖　商品名 Ficoll，相对分子质量约 400000，渗透压低，但因分子质量大而黏滞度高，为此常与泛影葡胺混合使用以降低黏滞度。主要用于分离各种细胞，如淋巴细胞，其他血细胞、成纤维细胞、肿瘤细胞、肝细胞等。

（3）氯化铯　是一种离子性介质，易溶于水，最高密度可达 1.91g/ml，分子量 168.4。由于它是重金属盐类，在离心时形成的梯度有较好的分辨率，被广泛地用于 DNA、质粒、病毒和脂蛋白的分离。

（4）卤化盐类　KBr 和 NaBr 可用于脂蛋白分离，KI 和 NaI 可用于 RNA 分离，其分辨率高于铯盐。NaCl 梯度也可用于分离脂蛋白，NaI 梯度可分离天然或变性的 DNA。

（5）Percoll　是商品名，它是一种 SiO_2 胶体外面包被了一层聚乙烯吡咯烷酮（PVP）。渗透压低，对生物样品影响不大，价格昂贵。Percoll 可用于细胞分离。

2. 等密度离心法

等密度离心法又分为预制梯度等密度离心法和自成梯度等密度离心法两种方式。

（1）预制梯度等密度离心法　该方法要求在离心前预先配制介质的密度梯度，常用的梯度介质为蔗糖、CsCl 和 Cs_2SO_4。待分离的样品一般也铺在梯度液顶上，其密度应小于顶部的梯度密度。如果夹在梯度液中间或管底部，应调节样品密度，以防未离心就上浮或沉淀。在进行离心时，由于不同种类的颗粒具有各自的浮力密度差，在离心场中，颗粒或向下沉降，或向上浮起，一直沿梯度移动到与它们密度恰好相等的位置上（即等密度点）形成区带，此即为等密度离心法。处于等密度点上的颗粒没有重量，区带的形状、位置均不受影响，体系处于动态平衡。此法因达到平衡所需时间较长，故需较长时间的离心。

（2）自成梯度等密度离心法　该方法是采用某些密度梯度介质经过离心后自身形成梯度。这种性能取决介质的本质，如蔗糖、甘油等梯度材料在离心后不会形成线性梯度，而 CsCl、Cs_2SO_4 等经长时间离心后可产生稳定的梯度。梯度介质（如 CsCl）由于离心力的作用逐渐形成管底部浓而管顶稀的密度梯度，与此同时原来分布均匀的颗粒也发生重新分布。需要离心分离的样品可和梯度介质先均匀混合，当管底介质的密度大于颗粒的密度即 $\rho_m > \rho_p$ 时，颗粒上浮；当管顶处 $\rho_p > \rho_m$ 时，则颗粒沉降；最后颗粒进入到一个它本身的密度位置即 $\rho_p = \rho_m$，颗粒不再移动。自成梯度等密度离心法总的离心时间应包括梯度介质形成密度梯度达到平衡的时间和样品颗粒进入等密度所需时间的总和，故所需时间较长，一般为十几小时至几十小时。离心时间与离心管内液柱长短有关，也与梯度材料的扩散系数有关。

三、离心操作注意事项

① 精密地平衡离心管和它们的内容物是十分重要的。超速离心机转头是镶置在一根很细的轴上，因此在两支离心管平衡时质量之差不得超过各个离心机说明书所规定的范围。制备性超速离心机两离心管平衡时质量之差不得超过 0.1g。

② 平衡后的一对离心管及其内容物（有时还可能包括离心管套筒）应对称放置，不能错位，以免因离心所产生的离心力不对称而损坏离心轴。

③ 当平衡的离心管及其内容物（或包括套筒）对称放置好后，盖好离心机盖子，打开电源，转动速度调节旋钮，逐步增加到所需要的离心速度，计算离心时间。

④ 当离心到达所需要的时间后，再转动速度调节旋钮减速。待离心机自然停止转动后取出离心管。关闭离心机电源，拔掉电源，清洁整理离心机、离心管、套筒、天平等。

⑤ 超速离心机带自动控制的操作系统，待选定离心速度、时间等参数后，按启动开关。

离心机即自动开始转动、记录时间，并可自动停机（取决于输入的控制信号）。在离心室被抽成真空前，离心机只作低速运转。

⑥ 超速离心时，应注意选择合适的离心管和转头。根据待离心液体性质及体积选择合适的离心管，装载液体时要按各种离心机具体操作说明进行。有的离心管无盖，液体不能装得过多，以防离心时甩出，造成转头生锈或被腐蚀。制备性超速离心机的离心管常常要求必须将液体装满，以免离心时塑料离心管上部凹陷变形。根据转速选择合适的转头，绝对不允许超过额定转速使用。转头在使用前应放冰箱内预冷。离心完后，转头应用温水洗涤并干燥，一般水洗就足够了，切勿将转头浸泡在去污剂中。转头在使用中应防止与其他物品碰撞，避免造成伤痕，如果发现有被腐蚀迹象，则不能再用，因为转头已经不安全了。

第五节　分光光度技术

分光光度技术（spectrophotometry）是根据物质对不同波长的光线具有选择性吸收，每种物质都具有其特异的吸收光谱而建立起来的一种定量、定性分析的技术，也称为吸收光谱法（absorption spectrometry）。其理论依据是朗伯-比尔定律。

分光光度技术是比色法的发展。比色法只限于在可见光区，分光光度技术则由可见光区扩展到紫外光区和红外光区。比色法用滤光片产生单色光，谱带宽度为 40～120nm，精度不高，而分光光度技术则采用棱镜或光栅产生单色光，其光谱带宽最大不超过 3～5nm，且具有较高的波长精度。本节仅讨论紫外及可见分光光度技术。

紫外及可见分光光度技术是一种分析技术。它不需要我们把欲分析的样品从混合物中分离开来，即可利用样品特殊的吸收峰或特殊的显色反应，直接进行定性、定量的分析。具有操作简单、灵敏度高、选择性强、定量分析的精密度和准确度都很高的优点。但是直接利用紫外光谱进行定性的能力相对较弱，通常还需与红外、色谱、质谱等技术结合才能作出可靠的定性鉴定。

一、基本原理——朗伯-比尔定律

利用分光光度技术来进行定量分析，测定物质含量的基本原理根据的是朗伯-比尔（Lamber-Beer）定律，这个定律讨论了有色溶液对单色光的吸收程度与溶液的浓度及液层厚度间的定量关系。此定律是由朗伯定律和比尔定律归纳而得。

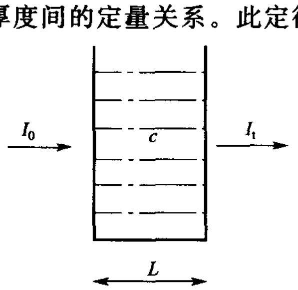


图 1-16　光吸收示意图

1. 朗伯定律

一束单色光通过溶液后，由于溶液吸收了一部分光能，光的强度就要减弱：若溶液浓度不变，则溶液的厚度愈大（即光在溶液中所经过的途径愈长），光的强度减低也愈显著，见图 1-16。

设光线通过溶液前的强度为 I_0（入射光的强度），通过液层厚为 L 溶液后，光的强度为 I_t（透过光的强度），则 I_t/I_0 表示透过光的强度是入射光强度的几分之几，称为透光度（transmittance），用 T 表示。透光度随溶液厚度的增加而减少，但实践证

明，透光度和溶液层厚度间并不存在简单的定量关系，只有透光度的负对数（$-\lg T$）才随着溶液厚度的增加而成正比例增加，即

$$-\lg T=-\lg \frac{I_t}{I_0}=\lg \frac{I_0}{I_t}\propto L$$

将上述比例写成等式，得到 $$\lg \frac{I_0}{I_t}=K_1 L$$

式中，$\lg \frac{I_0}{I_t}$称为吸光度（absorbance，A），又称为消光度（degree of extinction，E）或光密度（optical density，D），所以

$$A=K_1 L$$

式中，K_1 为比例系数，其值取决于入射光的波长、溶液的性质和浓度以及溶液的温度等。

上式表明，当溶液的浓度不变时，吸光度与溶液液层的厚度成正比，这就是朗伯定律。

2. 比尔定律

当一束单色光通过有色溶液后，溶液液层的厚度不变而浓度不同时，溶液浓度愈大，则透射光的强度愈弱，其定量关系如下：

$$\lg \frac{I_0}{I_t}=K_2 c$$

$$A=K_2 c$$

式中，c 为有色物质溶液的浓度；K_2 为比例系数，其值取决于入射光的波长、溶液的性质和液层的厚度，以及溶液的温度等。

上式表明，当溶液液层的厚度不变时，吸光度与溶液的浓度成正比，这就是比尔定律。

3. 朗伯-比尔定律

如果同时考虑吸收层的厚度和溶液浓度对光吸收的影响，将朗伯定律和比尔定律合并起来，得到

$$\lg \frac{I_0}{I_t}=KLc$$

$$A=KLc$$

这就是朗伯-比尔定律，它表示在单色光的条件下，溶液的吸光度与溶液浓度和液层厚度的乘积成正比。式中 K 是比例常数，式中若 L 用厘米表示，c 用质量浓度或摩尔浓度表示，则比例常数 K 称为吸光系数或消光系数。

二、分光光度技术的应用

分光光度分析是获得物质光吸收特性及定量、定性信息的重要手段。其应用可归为两类：一是在固定波长下测定物质溶液的吸光度，进行定量分析；二是在一定的波长范围内，测定波长对光密度的函数——吸收光谱，进行定性分析。

（一）定量分析

紫外可见分光光度技术最主要的应用是在定量分析方面。按照 Lamber-Beer 定律，溶液中溶质的吸光度正比于溶质的浓度。在一特定波长单色光下测出溶液的吸光度，即可计算出溶液的浓度。在使用紫外可见分光光度技术进行定量分析时，一般采用下列方法。

1. 标准曲线（工作曲线）法

配制含有与被测组分相同的一系列标准溶液，在被测组分的最大吸收波长（λ_{max}）下，

测定各标准溶液的 A 值，以 A 值为纵坐标，浓度为横坐标，绘出 A-c 曲线。将样品溶液在相同条件下测定 $A_{样}$ 值，从横坐标上找出与此 $A_{样}$ 值相对应的浓度，即为样品溶液的浓度。标准曲线的绘制应根据几次实验测得值的平均值进行绘制，这种方法对于大量样品分析或例行测定是比较方便的。在仪器和方法都固定不变的条件下，绘制好标准曲线后可多次使用，不必每次实验都再进行绘制。但如果以后测定的条件有所改变，如仪器经搬动后灵敏度有所改变、试剂用完后重新配制、温度有较大变化等，此时需对标准曲线进行校正或重新绘制。

在实际工作中，测定标准系列的吸光度 A 值后绘制 A-c 曲线时，常出现标准曲线只在一定浓度范围内呈直线关系，当浓度超过一定数值时，其吸光度的增加逐渐趋缓，不再遵守比尔定律。可以此浓度作为上限，重新调整标准系列的浓度并绘制标准曲线。但如果被测溶液浓度在曲线的直线范围内，也可直接进行定量测定。

标准曲线的制备不仅可用于样品的直接定量，而且在一定程度上还可检定所用的方法是否符合 Lamber-Beer 定律，对于实验技术人员（尤其是初始工作人员）来讲，也可检定其操作技术及实验条件等的控制是否正确无误。同时通过标准曲线的制备还可确定被测样品的最大浓度范围。

2. 对比法

分光光度法可用对比法测定溶液中物质的含量。它只需要利用空白调零后，分别测定标准溶液（浓度已知的溶液）和样品溶液（浓度待测定的溶液）的吸光度，然后进行比较。

$$\frac{A_x}{A_s}=\frac{Kc_xL}{Kc_sL}=\frac{c_x}{c_s},\text{即 } c_x=\frac{A_x}{A_s}\times c_s$$

式中，c_x 代表样品溶液的浓度；c_s 代表标准溶液的浓度；A_x 和 A_s 分别代表样品液和标准液所测得的吸光度值；只有 c_x 是未知的，可由上式计算得之。

此法适用 A-c 线性良好，且通过原点的情况，由于实际工作中所得到的 A-c 直线往往并非十分理想，为了减少误差，所用标准溶液的浓度应尽可能地与样品液的浓度相接近。

3. 计算法

在实际工作中，有时会遇上标准物缺乏或不易获得，此时可根据文献所提供的摩尔吸光率（ε）或在制备标准曲线时所求得的平均吸收系数，在相同的方法和条件下，直接测定样品在光径（液层厚度）为 10mm 时溶液的吸光度（A），按下式计算出样品的浓度（c）。

$$c=\frac{A}{\varepsilon}$$

（二）定性分析

使用分光光度计可以绘制吸收光谱曲线，方法是分别用各种波长不同的单色光分别通过某一浓度的溶液，测定此溶液对每一种单色光的吸光度，然后以波长为横坐标，以吸光度为纵坐标绘制吸光度-波长曲线，此曲线即吸收光谱曲线。中、高档次的分光光度计均可通过波长扫描，自动绘制出样品的吸收光谱曲线。

各种物质有它自己一定的吸收光谱曲线，因此用吸收光谱曲线图可以进行物质种类的鉴定。当一种未知物质的吸收光谱曲线和某一已知物质的吸收光谱曲线形状一样时，则很可能它们是同一物质。一定物质在不同浓度时，其吸收光谱曲线中峰值的大小不同，但形状相似，即吸收高峰和低谷的波长是不变的。

紫外光吸收是由不饱和结构造成的，含有双键的化合物能表现出吸收峰。紫外吸收光谱

比较简单，同一种物质的紫外吸收光谱应完全一致，但具有相同吸收光谱的化合物其结构不一定相同。除特殊情况外，单独依靠紫外吸收光谱不能决定一个未知物的结构，需要与红外、质谱等方法相配合。紫外吸收光谱分析主要用于已知物质的定量分析和纯度分析。若发现样品的吸收光谱存在有异常吸收峰，可以认为该异常吸收峰是由于样品中存在的杂质所致。

（王淳本　尹燕华）

第二章　蛋白质、核酸的提取与分离

蛋白质、核酸是重要的生物大分子，在生物体内具有重要作用，是分子生物学研究主要的对象。从生物体中提取与分离特定种类的蛋白质或核酸分子，通常是分子生物学研究的首要任务和工作基础。由于蛋白质与核酸的结构不同，理化性质也存在较大的差异，因此它们的提取与分离方法显著不同。然而，提取与分离蛋白质或核酸的基本步骤与原理相似，其一般流程是：首先采用适当的溶剂提取样品，使其进入溶液，随后用各种不同方法分离纯化，获得目标组分，最后经各种分析、鉴定方法确定终产物。

第一节　蛋白质的提取、分离纯化及定量

提取是将经过处理或破碎的细胞置于特定的溶液环境中，让被提取的生物大分子充分释放的过程。影响提取的因素主要是被提取物质在提取液中溶解度的大小及由固相扩散到液相的难易程度。某一物质在溶剂中溶解度大小与该物质的分子结构及溶剂理化性质有关，一般遵守“相似相溶”的原则。扩散作用对生物大分子的提取有重要影响。减小溶剂的黏度、增加搅拌强度及延长提取时间等，均可提高扩散速度，增加提取效果。提取过程应遵循“少量多次”的原则。对于所提取的蛋白质（或酶），必须做进一步的纯化工作，以除去大量与目标产物理化性质存在不同差别的杂质，已有多种纯化方法可供选择，实际工作中应根据纯化的目标和纯化产物的用途选择合适的纯化方案。

一、蛋白质（包括酶）的提取

一般而言，大部分蛋白质可以溶于水、稀盐、稀碱或稀酸溶液，少数与脂类结合的蛋白质则溶于乙醇、丙酮、丁醇等有机溶剂。蛋白质在不同溶剂中的溶解度差别，主要取决于蛋白质分子中极性基团与非极性基团的比例和这些基团的排列位置及偶极矩，因此采用不同溶剂和调整影响蛋白质溶解度的外界因素，如温度、pH、离子强度等，即可把所需的蛋白质和酶从细胞内复杂的组分中提取分离出来。

（一）水溶液提取法

稀盐缓冲系统的水溶液对蛋白质稳定性好、溶解度大，是提取蛋白质最常用的溶剂。加入的提取液的量要适当，加入量太少提取不完全，加的量太多，则不利于浓缩，通常用量是原材料体积的1～5倍，可一次提取或分次提取。提取时需要均匀地搅拌，以利于蛋白质的溶解。提取的温度视有效成分的性质而定。一方面，多数蛋白质的溶解度随着温度的升高而增大，因此，温度高利于溶解，缩短提取时间；但另一方面，温度升高会使蛋白质变性失活，因此，提取蛋白质时一般采用低温（5℃以下）操作。

为了避免蛋白质提取过程中的降解，可加入蛋白水解酶抑制剂（如二异丙基氟磷酸、碘

乙酸等)。下面着重讨论提取液的 pH 和盐浓度的选择。

1. pH

蛋白质是具有等电点的两性电解质，提取液的 pH 应选择在偏离等电点两侧的 pH 范围内。用稀酸或稀碱提取时，应防止过酸或过碱而引起蛋白质可解离基团发生变化，从而导致蛋白质构象的不可逆变化。如细胞色素 c 是一碱性蛋白质，常用稀酸提取；肌肉中甘油醛-3-磷酸脱氢酶是一酸性蛋白质，则用稀碱提取；植物组织中的一些酸性或碱性蛋白质常分别用 0.1%～0.2%的氢氧化钾水溶液或 1%碳酸钠溶液提取。在某些情况下，为了破坏目的蛋白质与其他杂质的静电结合，选择偏酸性（pH3～6）或偏碱性（pH10～11）的溶液提取，这样可以使离子键破坏而获得单一的蛋白成分。

2. 盐浓度

稀盐溶液可促进蛋白质的溶解，称为盐溶作用。同时，稀盐溶液因盐离子与蛋白质部分结合，具有保护蛋白质不易变性的优点，因此在提取液中加入少量 NaCl 等中性盐，一般以 0.15mol/L 浓度为宜。缓冲液常采用 0.02～0.05mol/L 磷酸盐或碳酸盐的等渗盐溶液。在某些情况下，也用到较高的盐浓度，如提取脱氧核糖核蛋白及膜蛋白。有时为了螯合某些金属离子和解离酶分子与其他分子的静电结合，选用柠檬酸缓冲液和焦磷酸钠缓冲液可获得较好效果。故稀盐溶液的浓度及缓冲液的组分的选择应根据不同对象及具体情况而定。总的来说，能溶于水溶液而与细胞颗粒结合较松的蛋白质或酶，在细胞破碎以后，只要选择适当的盐浓度和 pH，一般是不难提取的。

3. 温度

蛋白质和酶一般都不耐热，所以提取时通常要求低温（5℃以下）操作。只有对某些耐高温的蛋白质或酶（如胃蛋白酶、酵母醇脱氢酶及某些多肽激素）才在比较高的温度下提取，使更有利于和其他不耐热蛋白质的分离。

（二）有机溶剂提取法

一些和脂质结合比较牢固或分子中非极性侧链较多的蛋白质，难溶于水、稀盐溶液、稀酸或稀碱中，可用乙醇、丙酮或丁醇等有机溶剂提取，这些溶剂既有一定的亲水性，又有较强的亲脂性，是理想的提取液，但必须在低温下操作。

丁醇提取法对提取一些与脂质结合紧密的蛋白质和酶特别有利，一是因为丁醇亲脂性强，特别是溶解磷脂的能力强且不会引起蛋白质的变性失活；二是丁醇兼具亲水性，在 0℃时于水中有 10.5%的溶解度。它在水和脂分子间起着类似去污剂的桥梁作用，在取代蛋白质与脂质的结合位置后，由于丁醇与蛋白质极性基团结合的竞争力比脂质大，能阻止脱落了的脂质重新与蛋白质结合，使原来蛋白质在水中溶解度大大增加。另外，丁醇提取法的 pH 及温度选择范围较广，适用于动植物及微生物材料。

有些蛋白质和酶既溶于稀酸、稀碱，又能溶于一定比例的有机溶剂。在这样的情况下，采用稀的有机溶剂提取常常是为防止水解酶的破坏，并兼有除杂质和提高纯化效果的作用。如提取垂体前叶 ACTH 常采用酸性 70%丙酮，其设计原理是酸性可以促进 ACTH 的溶解，而 70%浓度丙酮对 ACTH 的溶解度没有影响，却可大幅度地抑制其他杂蛋白的溶出，对提高分离纯化的效果极为显著。

二、蛋白质的分离纯化

蛋白质和酶溶剂提取后进一步分离纯化常用方法有如下几种：

① 根据蛋白质溶解度不同进行分离的方法，包括分段盐析、等电点沉淀及有机溶剂沉淀；

② 利用色谱技术进行分离的方法，包括反向柱色谱、凝胶过滤色谱、离子交换色谱、亲和色谱等；

③ 利用电泳技术对蛋白质进行分离的方法；

④ 利用超速离心分离制备蛋白质；

⑤ 透析、超滤等膜分离技术。

各类蛋白质和酶从细胞中提取分离后，进一步纯化的方法选择及操作步骤繁简都不相同，很难作统一规定。但对于同一类的蛋白质，在提取分离上仍有许多共同点。

（一）根据蛋白质溶解度不同进行分离的方法

1. 蛋白质的盐析

中性盐对蛋白质的溶解度有显著影响，一般在低盐浓度下随着盐浓度升高，蛋白质的溶解度增加，称为盐溶；当盐浓度继续升高时，蛋白质的溶解度不同程度下降并先后析出，这种现象叫盐析。将大量盐加到蛋白质溶液中，高浓度的盐离子（如硫酸铵的 SO_4^{2-} 和 NH_4^+）有很强的水化作用，可夺取蛋白质分子的水化层，使之“失水”，于是蛋白质胶粒凝结并沉淀析出。盐析时若溶液 pH 在蛋白质等电点则效果更好。由于各种蛋白质分子颗粒大小、亲水程度不同，故盐析所需的盐浓度也不一样，因此调节混合蛋白质溶液中的中性盐浓度可使各种蛋白质分段沉淀。

影响盐析的因素有以下几个方面。①温度，除对温度敏感的蛋白质在低温（4℃）操作外，一般可在室温中进行。一般，温度低蛋白质溶解度降低。但有的蛋白质（如血红蛋白、肌红蛋白、清蛋白）在较高的温度（25℃）比 0℃时溶解度低，更容易盐析。②pH，大多数蛋白质在等电点时，在浓盐溶液中的溶解度最低。③蛋白质浓度，蛋白质浓度高时，欲分离的蛋白质常常夹杂着其他蛋白质一起沉淀出来（共沉现象）。因此在盐析前血清要加等量生理盐水稀释，使蛋白质含量在 25～30g/L。

蛋白质盐析常用的中性盐主要有硫酸铵、硫酸镁、硫酸钠、氯化钠、磷酸钠等。其中应用最多的是硫酸铵，它的优点是溶解度的温度系数小而溶解度大（25℃时饱和溶液为 4.1mol/L，即 541.2g/L；0℃时饱和溶解度为 3.9mol/L，即 514.8g/L），在这一溶解度范围内，许多蛋白质和酶都可以盐析出来；另外，硫酸铵分段盐析效果也比其他盐好，不易引起蛋白质变性。硫酸铵溶液的 pH 常在 4.5～5.5 之间，当用其他 pH 进行盐析时，需要硫酸或氨水调节。

硫酸铵浓度常以百分浓度表示。以 $X\%$ 表示所要配制的溶液浓度，$X_0\%$ 表示开始的溶液的浓度，所需的硫酸铵克数按以下公式计算：

$$硫酸铵克数=515(X-X_0)/(100-0.27X)$$

蛋白质在用盐析沉淀分离后，需要将蛋白质中的盐除去，常用的办法是透析，即把蛋白质溶液装入透析袋内，用缓冲液进行透析，并不断地更换缓冲液。因透析所需时间较长，所以最好在低温中进行。此外，也可用葡聚糖凝胶 G-25 或 G-50 过柱的办法除盐，所用的时间比较短。

2. 等电点沉淀法

等电点沉淀法主要利用两性电解质分子在电中性时溶解度最低，而不同的两性电解质具有不同的等电点而进行分离的一种方法。蛋白质、酶、核酸等生物大分子都是两性电解质。

在等电点的 pH 环境中，蛋白质颗粒之间的静电斥力最小，因而溶解度也最小，再加上其他沉淀因素，则很易沉淀析出。如不加其他沉淀因素，带有水膜的大分子（如蛋白质等）仍有一定溶解度，不易沉淀析出。由于许多蛋白质的等电点十分接近，单独使用此法效果不理想，分辨率也差，常用于提取液中去除杂蛋白。利用等电点除杂蛋白方法须事先了解所制备的物质对酸碱的稳定性，不然，盲目使用是很危险的。等电点法常和盐析法、有机溶剂法或和其他沉淀剂一起使用，以提高其沉淀能力。

3. 低温有机溶剂沉淀法

用与水可混溶的有机溶剂，如甲醇、乙醇或丙酮，可使多数蛋白质溶解度降低并析出。此法分离效率比盐析高，但蛋白质较易变性，应在低温下进行。

（二）利用色谱技术对蛋白质进行分离纯化

色谱技术是生物大分子研究中必不可少的工具，已被普遍认为是当前分离生化样品最有效的手段。不管是哪一种色谱方法，其最基本的特征是具有一个固定相和一个流动相，各种物质得以分离的最基本原因就在于它们在这两个相间具有不同的分配系数。当两相作相对运动时，这些物质在两相间进行反复多次的分配，从而得到分离。色谱技术有气相色谱法和液相色谱法两种。气相色谱以气体作流动相，液相色谱以液体作流动相。在生物大分子分离中，通常是以液相色谱技术为主。同其他分离技术相比，色谱方法具有以下优点：

① 分离效率高，选择性好，适合用于多种多元复杂混合物的分离；

② 既是精细的分离纯化方法，也是快速、灵敏、准确、简单的分析检测手段；

③ 可以处理从痕量至较大规模制备的样品量，可以实现在线检测和自动化操作。

表 2-1 列出了分离蛋白质常用的各种色谱技术。

表 2-1　用于蛋白质纯化的液相色谱技术

分离依据	分离方法（色谱类型）	特　点	用　途
分子大小	凝胶过滤	分级分离，分辨率适中，适于脱盐、分级分离时流速较慢（>8h 循环）的情况	大规模纯化的最后一步。用于去除微量杂质及聚合体脱盐，可用于任何阶段
电荷	离子交换	分辨率通常较高，流速较快（合适的填料），容量很大，样品体积不受限	最适于早期纯化，即大体积样品及蛋白纯度较低时使用
等电点	色谱聚焦	分辨率较高，流速快，容量很大，样品体积不受限	纯化最后阶段
疏水特性	疏水作用	分辨率较高，流速快，容量大，样品体积不受限	适于任何阶段，特别适于离子强度较高时，如沉淀、离子交换和亲和色谱后
	反相	分辨率较高，流速很快，容量大	适于最后阶段，特别适于相对分子质量较小的肽
生物亲和性	亲和	分辨率极高，流速大，容量大，体积不限	适于任何阶段，特别是样品浓度小、杂质含量多时使用

1. 凝胶过滤色谱

凝胶过滤色谱也称为体积排阻色谱，也可称为分子筛色谱。凝胶过滤色谱的机理是分子筛效应，如同过筛那样，它可以把物质按分子大小不同进行分离，但这种“过筛”与普通的过筛不一样。凝胶色谱柱中装填的是许多直径小于 1mm 的凝胶颗粒，在凝胶颗粒内部又是由三维多孔网状结构构成。在洗脱过程中，当分子大小不一的混合物样品加到凝胶床面后，大分子物质因其分子直径大于凝胶网孔而不能进入凝胶颗粒内部，只能沿着凝胶颗粒间的空隙随溶剂流动，受到的阻滞作用小，流程短而移动速度快，先流出色谱柱；分子量小的组分

则由于其分子直径小于网状结构的孔径而进入凝胶颗粒内部，受到网孔的阻滞作用，流程长而移动速度慢，比大分子物质后流出色谱柱，从而使分子大小不同的混合物得到分离。

目前所用的凝胶过滤的载体可以根据它们的孔径分为很多种不同的规格。每种规格都有其特定的分级范围，一旦超出分级范围，不论是其上限还是下限，即使分子量有大有小，但载体的排阻情况基本相同，就不能分离。为此在使用这种色谱时，要根据所要纯化的样品的分子量选择所用的凝胶过滤的基质。这种色谱方法，主要是应用于水溶性分子；但是也有应用于疏水性分子专用的凝胶过滤基质，只是种类很少。

凝胶过滤色谱最大的优点是整个色谱处理过程非常“温和”，可以尽可能地保护样品的生物活性。而不足之处是整个过程中样品非但不能浓缩，反而不断地在稀释，并且上样量也不能太大，否则分辨率下降，为此进行凝胶过滤的样品的浓度越大越好。

(1) 凝胶色谱的特点

① 由于凝胶色谱是按分子大小不同而分离，洗脱剂种类不影响其分离度，流动相的选择主要考虑对样品的溶解性及生物样品活性的保护。

② 凝胶色谱中无需改变洗脱液成分或种类。一次装柱后凝胶可反复使用，而且每次洗脱过程即是凝胶的再生过程。

③ 实验具有高度的重复性，回收样品几乎可达100%，既可用于大样品制备，亦可用于小样品的分析，所以应用广泛。

(2) 凝胶色谱法的应用

① 用于物质的分离提纯。由于凝胶的生物相容性较好，而且具有高质量的生物学活性回收率，所以凝胶色谱法主要用以分离和纯化分子大小不同的生物大分子物质。凝胶色谱法已广泛应用于酶、蛋白质、氨基酸、核酸、核苷酸、多糖、激素、抗生素、生物碱等物质的分离提纯，尤其与其他技术配合应用效果更为显著。由于凝胶色谱分离纯化蛋白质的量比较小，所以常常放在纯化工艺的最后一步。

② 作为脱盐工具。高分子（如蛋白质、核酸、多糖等）溶液中的盐类杂质可以借助凝胶色谱法将其除去，这一操作称为脱盐。凝胶色谱法脱盐速度快而完全，而且蛋白质、酶类等成分在分离过程不易变性。葡聚糖凝胶 Sephadex G-25 因流动阻力小、交联度适宜，常用于蛋白质溶液的脱盐。进行蛋白质样品的脱盐，应注意蛋白质溶液在去除电解质后，因溶解度显著下降以致成为沉淀物析出，从而被吸附在柱上不能洗脱下来。解决办法是，先使用有挥发性的盐类如甲酸铵、乙酸铵等缓冲液洗脱，取得洗脱液后再用冷冻干燥法除去挥发性盐类。

③ 用于去除热原物质。热原物质是微生物产生的某些使人发热的物质，如某些多糖蛋白质复合物等，用凝胶处理去离子水，可以得到适于制备注射剂的无热原水。

④ 蛋白质和多肽的相对分子质量测定。对不同的样品选用不同的凝胶过滤基质，然后用不同已知相对分子质量的样品（或几种样品的混合物）上柱，分别测定它们的洗脱体积（V_e）或洗脱时间（t），再由未知样品的 V_e 或 t 从上述关系中推测出未知样品的相对分子质量。在凝胶过滤所用的溶液系统中如不加变性剂等破坏四级结构，所测得的蛋白质样品的相对分子质量应是全蛋白的相对分子质量，一旦实验的条件能破坏四级结构，凝胶过滤所得的蛋白质的相对分子质量则是亚基的相对分子质量。一个蛋白质如由数种相对分子质量不同的亚基组装成，则在解聚条件下进行凝胶过滤就可能得到几个不同的蛋白质峰。

⑤ 蛋白质溶液的浓缩。干燥的葡聚胶颗粒内部有总值等于 V_i 的孔隙容积。当把干燥的

凝胶颗粒投入稀的高分子溶液时，水分和低分子物质就会进入凝胶粒子内部的孔隙直到充满为止，而高分子物质则排阻于凝胶颗粒之外，因此经几十分钟后，通过离心或过滤，就可以分离出膨胀的凝胶颗粒，得到浓缩了的高分子物质溶液，其中离子强度和 pH 都保持不变。这种浓缩方法特别适用于不稳定的蛋白质溶液的浓缩。

2. 离子交换色谱

离子交换色谱也是蛋白质分离纯化中使用得最广的色谱技术之一。其原理是利用样品的电荷和色谱载体的电荷间的相互作用，进而达到分离纯化的目的。离子交换剂有阳离子交换剂（如羧甲基纤维素，CM-纤维素）和阴离子交换剂（如二乙氨基乙基纤维素，DEAE-纤维素），当被分离的蛋白质溶液流经离子交换色谱柱时，带有与离子交换剂相反电荷的蛋白质被吸附在离子交换剂上，随后用改变 pH 或离子强度的办法将吸附的蛋白质洗脱下来。

在用离子交换树脂进行生物分子分离时，pH 是很重要的因素。对蛋白质或多肽类两性电解质而言，在不同的 pH 时带的电荷不同，这样，不同的 pH 和不同的离子强度下，蛋白质或多肽在不同的离子交换树脂上的吸附程度也有不同。如果能选择适当的条件，对蛋白质的分离纯化能起到事半功倍的效果。现以血浆蛋白中的两个蛋白质的分离纯化为例。一是免疫球蛋白 G（IgG），它的等电点较高，故在较高 pH 和较低离子强度时，就不能吸附在以此条件的阴离子交换树脂上。因此，在硫酸铵分级，收集 0～40%饱和度的沉淀后，用 pH 8 左右的缓冲液（含约 50mmol/L 的 NaCl）溶解并透析，然后上样于用同样缓冲液平衡的 DEAE 离子交换树脂，结果只有 IgG 没有吸附，而这一级其他的蛋白几乎全吸附在柱上。α_1-酸性糖蛋白是一个等电点很低的蛋白质，因此，经硫酸铵分级的血浆蛋白，40%饱和度的上清部分对 pH4.0 缓冲液（含约 50mmol/L 的 NaCl）透析，也上 DEAE 离子交换树脂，只是这次树脂是用 pH4.0 含 50mmol/L NaCl 的缓冲液平衡，结果绝大多数蛋白质不吸附在柱上，仅 α_1-酸性糖蛋白吸在柱上，提高洗脱液中的盐浓度，就能得到较纯的样品。

离子交换色谱也有浓缩样品的作用。这种色谱基本上是用于亲水的物质。由离子交换色谱所选用的树脂的类型以及样品洗脱的条件，还可以对样品的带电行为提供一些信息。此外，离子交换树脂价格低廉，因此广泛用于蛋白质等生物活性物质的制备。

另有一种色谱聚焦技术（chromatofocusing）。有一类特殊的商品化的离子交换树脂，Sigma 的产品目录上称为聚缓冲液交换剂（polybuffer exchanger）。先用某一 pH 的缓冲液平衡聚缓冲液交换剂，再用另一 pH 的缓冲液平衡，就能在这种离子交换剂上建立一个 pH 梯度。上样以后，用一种具有 pH 梯度的聚缓冲液洗脱树脂时，被吸附样品中的蛋白质就能按照它们的等电点次序被洗脱。在该过程中，同时产生了聚焦作用，致使各蛋白质区带变得很窄，样品也高度浓缩，整个分离过程也呈现出很高的分辨率。

3. 亲和色谱

亲和色谱是利用生物大分子与某些相对应的专一分子特异识别和可逆结合的特性而建立起来的一种分离生物大分子的色谱方法，也叫做生物亲和色谱或生物特异性亲和色谱。这种特异可逆结合的物质很多（见表 2-2），如抗原与抗体、底物与酶、激素与受体等，它们间的这种特异亲和能力又叫亲和力。如果将生物分子间的亲和力视为特殊的吸附作用，那么亲和色谱也能归属到吸附色谱一大类。

亲和色谱的最大特点是高效和简便。一旦固定化配体制备后，操作使用非常方便，通常情况只要在上样后洗去杂质，就可将所要的物质解吸下来，而且所得的物质已有较高的纯度。这样的方法可以从含量极低的原料中经一次亲和色谱就能分得所要的物质，回收率也较

表 2-2 亲和色谱所用体系

常用配体	分离对象	常用配体	分离对象
酶	底物、抑制剂、辅酶	核酸	互补碱基序列、组蛋白、核酸聚合酶、结合蛋白
抗体	抗原、病毒细胞	激素	受体、载体蛋白
外源凝激素	多糖、糖蛋白	细胞	细胞表面特异蛋白、外源凝激素

高。以血浆中的维生素 B_{12} 结合蛋白为例，这种结合蛋白在血浆中的含量仅 5/10000000，如果应用常规的分离纯化方法，如盐析、离子交换色谱或电泳等，即使能得到，回收率也很低。采用固定化的维生素 B_{12}，很简便地就能纯化出这一结合蛋白。亲和色谱的过程也是高度浓缩的过程。亲和色谱还有另外两个优点，是其他的分离纯化方法所不能比的。一是可以从样品中除去大量的杂质，回收率较高地得到少量的物质；二是可以在活性物质中除去理化性质几乎完全相同的但已失活了的那些“杂质”，这一点对提高酶或其他分子的比活特别有用。

在亲和色谱中有三种具有较大通用性的技术，这就是免疫亲和色谱、凝集素亲和色谱和染料亲和色谱。将抗体偶联到载体所形成的免疫亲和柱可用于抗原的分离和纯化。凝集素是一大类能专一性地和糖类及糖复合物结合的蛋白质，固定化的凝集素则可用于糖蛋白的分离纯化。

(三) 利用电泳技术对蛋白质进行纯化分离

因蛋白质分子由氨基酸组成，而氨基酸带有可解离的氨基（$-NH_3^+$）和羧基（$-COO^-$），所以蛋白质是典型的两性电解质，在一定的 pH 条件下就会解离带电，带电的性质和多少取决于蛋白质分子的性质及溶液的 pH 值和离子强度。蛋白质在不同 pH 环境中带电性质和电荷数量不同，各种蛋白质在同一 pH 条件下，因相对分子质量和电荷数量不同，在电场中的迁移率不同而得以分开（详见第一章有关内容）。

常规聚丙烯酰胺凝胶电泳（PAGE）和各种特殊的聚丙烯酰胺凝胶电泳［SDS-聚丙烯酰胺凝胶电泳（SDS-PAGE）、梯度聚丙烯酰胺凝胶电泳（PG-PAGE）、等电聚焦聚丙烯酰胺凝胶电泳（IEF-PAGE）、双向凝胶电泳（2D-PAGE）］等电泳技术是常用的蛋白质分离分析技术，琼脂糖凝胶电泳则常用于脂蛋白的分离分析。电泳技术具有很高的分辨率、灵敏度和稳定性，常用于蛋白质的定性和纯度的分析，也可制备纯度极高的蛋白质纯品。一般常见常用的电泳设备大多是分析型的，能处理样品的量很少，仅为数十到数百微克，利用电泳法分离纯化一定量的蛋白质，往往需要采用制备型的电泳设备。

(四) 利用超速离心分离制备蛋白质

制备超离心技术就是在强大的离心力场下，依据物质的沉降系数、质量和形状不同，将混合物样品中各组分分离、提纯的一项技术。在生物化学、分子生物学以及细胞生物学的发展中，制备超离心技术起着很重要的作用。迄今，这项技术已广泛用于各种细胞器、病毒以及生物大分子的分离。

由于许多具有生物活性的生物大分子和细胞器很不稳定，在分离时要求温度低、操作条件缓和，不能应用一般的化学分离法，只能利用制备超离心机来进行大量分离制备。下面介绍一个用 NaBr 等密度梯度离心分离血浆脂蛋白的实例。

① 用不含 NaBr 的稀液（11.4g NaCl 和 0.1g EDTA，以蒸馏水定容至 1L，pH 7.8）和浓 NaBr 液（263g NaBr、11.4g NaCl 和 0.1g EDTA，以蒸馏水定容至 1L，调 pH 7.8）

通过梯度发生器，在 Tepy 70Ti 式转子的离心管中制备 30ml 密度为 1.003～1.2g/cm^3 的 NaBr 线性梯度。

② 每毫升血清样品中加入 0.32g NaBr，将密度调到 1.22g/cm^3。

③ 将 4ml 样品分层到梯度的底部，样品层下面再铺放 1.5ml 浓 NaBr 液（0.582g/ml）。导入样品层应注意排除气泡，以免扰乱梯度。

④ 采用慢加速方式运行，从静止加速到 3000r/min 需要 3min 左右，然后按正常加速。在 4℃，40000r/min 下离心 3h。

⑤ 离心时间到后采用慢减速方式停机（按正常减速到 3000r/min 后，进行慢减速，从 3000r/min 减速到静止约需要 3min 左右）。

⑥ 通过向上取代法取出梯度。通常最低密度脂蛋白在梯度顶部，较低密度脂蛋白在顶部以下 15～20ml 之间；而高密度脂蛋白与其他血清蛋白一起在原样品位置。

（五）膜分离技术在蛋白质分离纯化中的应用

迅速发展的膜分离法已成为物理化学分离法中重要组成部分。它在生化制备中用于浓缩及脱盐已占压倒优势，用于分离提纯的事例逐年增加。膜分离技术的原理主要是利用溶液中溶质分子的大小、形状、性质等的差别，对于各种薄膜表现出不同的可透过性而达到分离的目的。选择薄膜在膜分离法中很重要。薄膜的作用是有选择地让小分子通过，而把较大的分子挡住，所以又常把这种膜称为半透膜。透析是以分子扩散作用作为较小分子透过薄膜的动力，而超滤则是以膜的两边外加的流体静压差（高压力或离心力）作为较小分子透过薄膜的动力。

1. 透析

透析法的特点是半透膜两边都是液相，一边是样品液，另一边是纯净溶剂（水或缓冲液）。样品液中不可透析的大分子被截留于膜内，可透析的小分子经扩散作用不断透出膜外，直到两边浓度达到平衡。透析法多用于制备及提纯生物大分子时除去或更换小分子物质、脱盐和改变溶剂成分。透析膜除动物膜外都是纤维素或纤维素衍生物制成的。商品透析膜大多为管状，各公司采用的大小规格又不相同，购买时需注意。表 2-3 列出 Union Carbide 各种型号透析管的渗透范围。

表 2-3　Union Carbide 各种型号透析管的渗透范围

型　号	近似膨胀直径/cm	可透过物质的分子量	不能透过的分子量
8 号透析管	0.62	5732	20000
18 号透析管	1.40	3300	5732
20 号透析管	1.55	30000	45000
27 号透析管	2.10	5732	20000
36 号透析管	2.80	20000	—

商品透析管膜常涂甘油以防破裂，并含有极其微量硫化物、重金属和一些具有紫外吸收的杂质。它们对蛋白质和其他生物活性物质有害，用前必须除去。有人建议先用 50％乙醇慢慢煮沸 1h，再分别用 50％乙醇、0.01mol/L 碳酸氢钠溶液、0.0001mol/L EDTA 溶液依次洗涤，最后用蒸馏水浸洗 3 次，基本可除杂质。实验证明，50％乙醇处理对除去具有紫外吸收杂质特别有效。已处理好的管膜如果不用，可贮存于 4℃蒸馏水中。如需长期贮存，可加少量叠氮化钠、氯仿以防细菌侵蚀；使用时先用蒸馏水充分漂洗，再用透析液漂洗即可使用。

透析方法很简单，可将已处理及检查过的透析袋用夹子或尼龙丝扎紧底端，然后将待透析的样品液（1～100ml）从管口倒入袋内，但不能装满，常留一半左右的空间，以防膜外溶剂大量透入袋内时将袋胀裂。装透析液后，即紧扎袋口，悬于装有大量纯净溶剂（水或缓冲液）的大容器内（量筒或玻璃缸）进行透析。小分子（水或盐）可从透析膜内透出，直到膜内外浓度相等。如果加上搅拌以及定期（或连续）换上新鲜溶剂，可大大提高透析效果。如需冷却，可在冰箱内进行。

2. 超滤

超滤是加压膜分离技术之一，操作简便，成本低廉，分离效率高，且不引起温度、离子状态及相的变化，近来广泛用于生物制品浓缩、脱盐及分级分离工序。超滤原理和一般过滤一样，主要依赖于被分离物质分子量的大小、形状和性质不同，在一定的压力差下，使小分子能够通过具有一定孔径的特制薄膜，限额以上的大分子被膜阻留，使不同大小的分子得以分离。

要在膜的两边产生压差，可在滤液一边加正压或在超滤液一边产生负压。因为抽真空产生的负压差不超过 1×10^5 Pa（一个标准大气压），超滤效果差，所以较少采用；现大多利用液压泵产生的压力进行超滤。还有许多小型超滤装置制作成离心管的规格和样式，放置于离心机的转头中，利用离心机高速旋转时强大的离心力场产生极高的压力进行超滤。

蛋白质、酶等大分子用一次超滤法很难达到分离纯化的目的，需用2～3级超滤装置才能收到较好效果。串联超滤装置是根据截留不同大小的分子来区分的。第一级装置选用的膜分子量截留值大于第二级，第二级用的膜分子量截留值大于第三级，其余依此类推。然后用色谱分析法测定各级超滤装置的滤出液及保留液中的成分，即可知各级的分离效果及组分的纯度。

（六）重组蛋白的分离纯化

从天然组织原材料中分离纯化蛋白质与从各种不同的表达细胞株中分离纯化相应的重组蛋白质有很大不同，相对来讲，重组蛋白质的分离纯化要比相应的天然蛋白质容易，这主要是因为表达系统中重组蛋白质含量较高，而杂蛋白相对较少，尤其是应用细菌和酵母作为表达系统时更是如此。此外，应用分泌细胞株表达重组蛋白质时所含杂蛋白也不多。

1. 分离纯化时必须考虑的因素

这些影响因素包括工程菌的种类或细胞类型、代谢特性、表达方式（胞内、胞外、包含体）、表达产物在细胞内所处的位置、表达产物和副产物种类、代谢物种类、产物类似物和能降解表达产物的酶类等，同时还必须考虑表达产物的重要性质，包括物理、化学与生物学特性，如化学组成、分子量、等电点、电荷分布和密度、溶解度、稳定性、疏水性、扩散性、分配系数、吸附性能、生物学活性、亲和性、配基种类和表面活性等，这些都是分离纯化的重要依据，必须充分考虑。尤其应该考虑到重组蛋白质都是具有生物活性的大分子，稳定性差，对pH、温度、金属离子、有机溶剂、剪切力、表面张力等十分敏感，容易变性失活，也容易被蛋白酶所降解，因此在选择、确定纯化方法时要十分注意对目的蛋白质的保护。

2. 分离纯化的一般过程

重组蛋白质的分离纯化一般包括细胞抽提与重组蛋白质的回收、浓缩与预处理、柱层析分离纯化等步骤。分离纯化过程一般在低温条件下进行，条件要温和，以保持重组蛋白质的生物活性；分离纯化方法的选择性要好，能从复杂的混合物中有效地将目的产物分离出来，

达到较高纯化倍数和回收率；纯化步骤之间要能直接衔接，以节约时间、减少损失、提高效率。

（1）细胞抽提与重组蛋白质的回收　在理想的情况下，重组蛋白质表达后经过翻译后修饰分泌到细胞培养液中，然后从培养液中直接回收重组蛋白质。例如，在动物细胞和酵母细胞表达系统中，通过离心或过滤就可以方便地将培养液与表达细胞分离。

但是，一些蛋白质，特别是在细菌中表达的蛋白质，表达量很高，以一种称为“包含体”的形式在细菌细胞中出现。它是一种不溶性的结构致密的蛋白质聚集体。

分泌型重组蛋白质的抽提和分离纯化的一般步骤为：经离心或过滤获得培养上清，然后用超滤或沉淀方法进行样品的浓缩，再经柱层析法获得纯化后的重组蛋白质。而包含体重组蛋白质的抽提和分离纯化的一般步骤为：破碎培养细胞，离心获得含重组蛋白质的沉淀，然后将沉淀溶解，经柱层析进行蛋白质的纯化，最后进行重组蛋白质的复性。

包含体中的重组蛋白质以变性的形式存在，但纯度很高。有些情况下，包含体也含有一些宿主细胞的生物成分。如果重组蛋白质不是以分泌形式表达，就需要通过各种方法将细胞破碎，包括物理、化学或生物学方法，使重组蛋白质释放出来。之后，可以用离心方法使包含体沉淀。细菌细胞壁的破碎要小心，使得细胞蛋白质留在上清，通过低速离心使其与包含体分离。在沉淀中重组蛋白质通常可以占到包含体蛋白质的40%到60%以上。如果发现重组蛋白质部分在包含体中，部分以可溶形式存在，可以加入酸或有机溶剂以使溶解的重组蛋白质沉淀。

（2）包含体的溶解和重组蛋白质的复性　包含体的溶解需要变性剂（如6～8mol/L的盐酸胍）或去垢剂（如Triton或SDS），加入巯基化合物，使二硫键还原，并将蛋白质转变为完全随机的结构形式。这时，溶解的包含体蛋白质获得完全变性，即除一级结构和共价键保留外，所有的氢键和疏水键都被破坏，疏水侧链完全暴露。接下来要通过蛋白质的纯化，使蛋白质易于复性，即完全恢复到具有生物活性的天然构象状态。在溶解的包含体中，重组蛋白质的纯度相对较高，因此，这时凝胶过滤是比较好的选择。如果变性剂是离子性的（如盐酸胍），需要用非离子性的变性剂（如尿素）进行透析，从而在接下来的纯化中应用离子交换或中性的吸附层析方法。一般都需要加入最低剂量的变性剂使重组蛋白质维持在溶解状态。在重组蛋白质的复性过程中，最好加入一定剂量的氧化型和还原型巯基化合物的混合物。

蛋白质复性是一个非常复杂的过程。一般可以采用二种方法。一种是稀释溶液，使溶液中的变性剂浓度逐步降低，蛋白质开始复性。该法简单易行，但有时操作体积过大，同时也降低了蛋白质的浓度。另一种方法是通过透析、超滤或电渗析缓慢除去变性剂，使蛋白质获得复性；此法溶液总体积和蛋白质浓度变动不大，但耗时较长，有时会形成蛋白质沉淀。复性通常是一个经验性的步骤，需要对变性剂、还原剂以及去除这些试剂的条件进行摸索。一般来说，重组蛋白质的分子量越小，二硫键越少，复性成功的可能性就越大。重组蛋白质如果含有两个以上的二硫键时，很有可能产生肽链内和肽链间二硫键的错配。为了形成正确的二硫键，常在包含体蛋白质复性时加入还原型谷胱甘肽和氧化型谷胱甘肽等氧化还原缓冲液。并且，为了抑制蛋白质聚集，防止分子间错误的二硫键形成，常使变性剂保持在1～2mol/L之间。当分子内二硫键少时，可不用氧化还原缓冲液，而通过搅拌，依靠空气的氧化作用，恢复表达蛋白的正确构型。

（3）分泌蛋白质的浓缩　如果表达系统将已折叠和翻译后加工的重组蛋白质分泌到细胞

培养液中，必须将大量的培养液先进行浓缩后，才能进行后续的纯化步骤。浓缩可采用超滤、沉淀、吸附/洗脱层析等方法。

当细胞抽提的体积浓缩到一定程度后，开始真正的纯化，目前主要依赖柱层析方法。对于完全新的重组蛋白质来讲，没有任何固定的纯化模式。纯化重组蛋白质的步骤与纯化相应的天然蛋白质的步骤也许相似，也许差别很大。但是，纯化天然蛋白质的方法对纯化相应的重组蛋白质具有重要的借鉴作用。

（七）分离制备蛋白质的一个实例

蛋白质在组织或细胞中是以复杂的混合物形式存在，每种类型的细胞都含有上千种不同的蛋白质，因此蛋白质的分离、提纯和鉴定是生物化学中的重要部分，至今还没有一个单独或一套现成的方法能把任何一种蛋白质从复杂的混合蛋白质中提取出来，因此往往采取几种方法联合使用。下面以制备鸡卵类黏蛋白为例进行介绍。

鸡卵类黏蛋白（简称 CHOM）是由鸡卵清中制得的一种糖蛋白，它具有强烈的抑制胰蛋白酶的作用，常用于胰蛋白酶的酶学性质研究。也可将其制成吸附亲和剂，通过亲和色谱技术有效地分离与纯化胰蛋白酶。

鸡卵类黏蛋白至今还未能获得单一组分的制品，在电泳行为上常呈不均一性。目前至少已获得 4 种不同组分，它们在抑制胰蛋白酶的性能上和氨基酸组成上没有多大区别，但是在糖蛋白的糖部分（主要为 D-甘露糖、D-半乳糖、葡萄糖胺和唾液酸）的含量上则有差别。它们的等电点有一定的范围，大致在 pH3.9～4.5；相对分子质量 28000。卵类黏蛋白抑制胰蛋白酶的物质的量的比为 1∶1，因此高纯度的卵类黏蛋白 1μg 能抑制相当于 0.86μg 的胰蛋白酶（比活力为 12000BAEE 单位/mg 蛋白）。不同来源的卵类黏蛋白末端氨基酸有很大差异，鸡卵类黏蛋白的 N 末端为丙氨酸，C 末端为苯丙氨酸。

卵类黏蛋白在中性或酸性溶液中对热和高浓度的脲都是相当稳定的，而在碱性溶液中比较不稳定，尤其当温度较高时易迅速失活。鸡卵类黏蛋白除对牛和猪的胰蛋白酶有强烈的抑制作用外，对枯草杆菌蛋白酶也有一定程度的抑制作用，但对胰凝乳蛋白酶无抑制作用，另外对人的胰蛋白酶也无明显的抑制作用。

本实验主要参照 Kassell 的方法，鸡蛋清先经三氯乙酸（TCA）溶液处理，除去沉淀物，然后经丙酮分级沉淀获得鸡卵类黏蛋白的粗品，再经 DEAE-纤维素（二乙基氨基乙基-纤维素）柱色谱纯化而得鸡卵类黏蛋白的合格产品。

1. 鸡卵类黏蛋白的提取

取 50ml 鸡卵清，加入等体积的 10%(pH 1.15) 三氯乙酸溶液（缓缓加入以防局部过酸出现块状物），这时出现大量的类似于酸奶状的白色沉淀，轻轻地搅拌均匀。在酸度计上检查提取液的 pH，此时提取液的最终 pH 大约是 3.5，若该溶液的 pH 偏离 pH3.5±0.2，则要用 5mol/L 氢氧化钠或 5mol/L 盐酸将 pH 调回到 pH3.5±0.2 的范围以内。由于提取液非常稠，在调节 pH 时要防止局部过酸或过碱。pH 调好、稳定后，在室温下静置 4h 以上。待卵清清蛋白完全沉淀后，以 3000r/min 离心 10min。弃去沉淀物，上清液用滤纸过滤，以除去上清液中的脂类物质及其他不溶物。收集滤液转移到一个 500ml 的烧杯内。检查滤液的 pH 是否为 pH3.5，否则要调回到 pH3.5。放置冰浴或冰箱内冷却片刻，缓慢加入 3 倍体积预冷的丙酮。用玻璃棒轻轻搅匀，用塑料薄膜盖好封严，防止丙酮挥发，在冰浴里放置 2～4h 左右，待鸡卵类黏蛋白完全沉淀后，小心吸出一部分上清液，剩余的部分全部转移到 50ml 带盖的离心杯里，盖上盖经平衡后以 3000r/min 离心 15min，除去上清，离心管底部

的沉淀物放在真空干燥器内，抽气除去残留丙酮。待沉淀物由白色转变为透明黏稠物后停止抽气。加入 20ml 蒸馏水溶解，若溶解液浑浊则用滤纸滤去不溶物。滤液经 Sephadex G-25 柱色谱法脱盐或者经透析除盐。Sephadex G-25 柱色谱法脱盐的操作如下。

称取 30g Sephadex G-25，用 500ml 0.02mol/L（pH6.5）的磷酸盐缓冲液在 100℃热溶胀 2h 或者室温下溶胀 24h。脱气后装柱（35mm×300mm），柱床体积约 150ml。用约 2 倍体积的 0.02mol/L（pH6.5）磷酸盐缓冲液平衡。流出液在核酸蛋白质检测仪上绘出稳定的基线或经紫外分光光度计测定光吸收，其 A_{280nm}小于 0.02 即可使用。

取 20ml 鸡卵类黏蛋白提取液上柱（上样量不超过柱床体积的 1/6）。用同样的缓冲液洗脱，收集第一峰，在蛋白峰完全流出后盐才开始流出，盐通常在 280nm 无明显光吸收，若绘出第二峰则是残留丙酮峰。丙酮和盐同时流出。洗脱曲线见图 2-1。

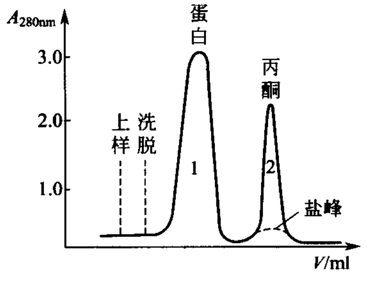


图 2-1　鸡卵类黏蛋白的 Sephadex G-25 柱色谱洗脱曲线

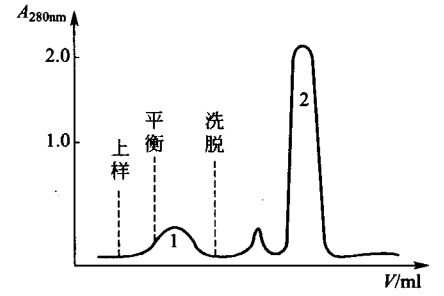


图 2-2　鸡卵类黏蛋白的 DEAE-纤维素离子交换柱色谱洗脱曲线

2. 鸡卵类黏蛋白的纯化

通过上述方法获得的鸡卵类黏蛋白仍含有少量的卵清清蛋白。由于鸡卵类黏蛋白与卵清清蛋白的等电点不同，当它们处在同一 pH 缓冲液中时，电离有所不同，因此，采用 DEAE-纤维素离子交换色谱可以将二者分开，从而达到分离纯化的目的。

称取 10g DEAE-纤维素粉（DE-32），以 150ml 0.5mol/L 氢氧化钠-0.5mol/L 氯化钠溶液浸泡 20min。转移到布氏漏斗内（内垫有 200 目的尼龙网）抽滤。用蒸馏水洗至 pH8.0 左右，抽干。然后移至一个 500ml 的烧杯内，用 150ml 0.5mol/L 盐酸溶液浸泡 20min，再转移到布氏漏斗内，抽滤，用蒸馏水洗至 pH6.0 左右，最后转移到烧杯内，用 150ml 0.2mol/L（pH6.5）磷酸盐缓冲液浸泡约 15min，经真空干燥器脱气后装柱。

取一支色谱柱（35mm×200mm）装入约 1/4 体积的 0.02mol/L（pH6.5）磷酸盐缓冲液，将处理过的 DEAE-纤维素装入柱内。以同一缓冲液平衡，流出液在核酸蛋白质检测仪上绘出稳定的基线或经紫外分光光度计测定光吸收，待 A_{280nm}值小于 0.02 即可。

将经 Sephadex G-25 脱盐后的鸡卵类黏蛋白溶液上柱吸附。以 0.02mol/L（pH6.5）的磷酸盐缓冲液平衡，待流出液在核酸蛋白质检测仪上绘出稳定的基线后，改用 0.3mol/L 氯化钠-0.02mol/L（pH6.5）的磷酸盐缓冲液洗脱，收集第二洗脱峰。柱色谱洗脱曲线见图 2-2。若在鸡卵类黏蛋白液中所含的鸡卵清清蛋白的量很少，在洗脱时可能出现一个小峰或出现不明显的峰形。在这种情况下可根据峰形大小和测定活性来确定鸡卵类黏蛋白洗脱峰的位置。

3. 鸡卵类黏蛋白纯品的制备

经 DEAE-纤维素离子交换柱色谱制得的鸡卵类黏蛋白溶液装入透析袋内，对蒸馏水进

行透析，间隔一段时间更换一次蒸馏水，直至经1%硝酸银检查无氯离子为止。

将透析过的鸡卵类黏蛋白溶液转移到烧杯，小心地用1mol/L盐酸调节至pH4.0，取出1ml，稀释5～10倍，测定鸡卵类黏蛋白的含量及活性。然后加入3倍体积的预冷丙酮，用塑料薄膜封严烧杯口，在冰浴里静置4h左右。待鸡卵类黏蛋白完全析出后，倾出上清液，沉淀部分转移到一个带盖的50ml的离心管内，于3000r/min离心15min，弃去上清，沉淀物放入真空干燥器内真空干燥，即可得到透明胶状物——鸡卵类黏蛋白（约200～300mg）。

4. 鸡卵类黏蛋白的活性测定

鸡卵类黏蛋白是胰蛋白酶的天然抑制剂，因此，它的活性可以用抑制胰蛋白酶的活力单位来表示，即抑制1个胰蛋白酶活力单位（BAEE单位）所需卵类黏蛋白量定为抑制剂的1个活力单位。将卵类黏蛋白和结晶胰蛋白酶按一定比例相互混合（卵类黏蛋白量不能超过胰蛋白酶量，一般以1∶2较合适，具体视卵类黏蛋白的纯度而异），加入适量的0.05mol/L（pH8.0）Tris-HCl缓冲液，在25℃保温10min左右，使酶与抑制剂充分结合。然后取适量混合液（相当于原胰蛋白酶10μg左右），按胰蛋白酶活力测定方法测出剩余酶活力，同时，取相同量的胰蛋白酶（未加抑制剂）按常规方法测出胰蛋白酶活力，由它减去剩余酶活力，即为被抑制的胰蛋白酶活力，也就是卵类黏蛋白的抑制活力。

取2个石英比色皿（带盖，光程为1cm），其中一个加入1.5ml 0.05mol/L（pH8.0）Tris-HCl缓冲液和1.5ml 2.0mmol/L BAEE底物溶液，混匀，在波长253nm处调整仪器零点。向另一个比色皿内加1.5ml 0.05mol/L（pH8.0）Tris-HCl缓冲液、10μl胰蛋白酶（约10μg胰蛋白酶）、10μl鸡卵类黏蛋白（约5～10μg鸡卵类黏蛋白）轻轻摇匀，在室温下（最好是在25℃恒温箱内）放置2min，使胰蛋白酶与鸡卵类黏蛋白充分结合。然后加入1.5ml 2mmol/L BAEE底物溶液，立即混匀并记时。于波长253nm处测定其光吸收递增值——$\Delta A_{253nm}/\text{min}$，每隔30s读数一次，使$\Delta A_{253nm}/\text{min}$值控制在0.05左右为宜，否则要根据$\Delta A_{253nm}/\text{min}$的大小适当减少或增加鸡卵类黏蛋白的量。

以反应时间t为横坐标，A_{253nm}为纵坐标作图，取直线部分的任一时间间隔与相应的光吸收变化值为$\Delta A_{253nm}/\text{min}$，用$A_2$表示。与此同时，以同样的方法（不加鸡卵类黏蛋白）测胰蛋白酶的活性，通过作图得到光吸收变化值$\Delta A_{253nm}/\text{min}$，用$A_1$表示。

鸡卵类黏蛋白的抑制活性及抑制比活性的计算公式如下：

$$\text{鸡卵类黏蛋白的抑制活性单位（BAEE单位）}=\frac{A_1-A_2}{0.001}$$

$$\text{鸡卵类黏蛋白抑制比活单位（BAEE单位/mg胰蛋白酶）}=\frac{(A_1-A_2)\times 1000}{I\times 0.001}$$

式中 A_1——胰蛋白酶$\Delta A_{253nm}/\text{min}$；

A_2——加入抑制剂后胰蛋白酶$\Delta A_{253nm}/\text{min}$；

I——测定时所用鸡卵类黏蛋白的量（μg）；

1000——转换系数，将抑制剂的质量单位由μg转换为mg；

0.001——光吸收每增加0.001定义为1个BAEE单位。

5. 鸡卵类黏蛋白的含量测定

配成一定浓度的鸡卵类黏蛋白溶液，在280nm处测定其光吸收。鸡卵类黏蛋白的消光系数是$A_{1cm}^{1\%}=4.13$，式中的1%为蛋白浓度（1g/100ml溶液），当蛋白浓度为1mg/ml时，其消光系数为0.413。

$$鸡卵类黏蛋白(mg/ml)=\frac{A_{280nm}\times 稀释倍数}{0.413}$$

6. 注意事项

① 提取鸡卵类黏蛋白时，提取液的最终 pH 大约是 3.5，若该溶液的 pH 偏离 pH3.5±0.2，则要用氢氧化钠或盐酸将 pH 调回到 pH3.5±0.2 的范围以内；由于提取液非常稠，在调节 pH 时要防止局部过酸或过碱。

② 纯化鸡卵类黏蛋白使用过的 DEAE-纤维素离子交换色谱柱，不要忘记用 0.02mol/L (pH6.5) 的磷酸盐缓冲液重新平衡，待流出液在核酸蛋白质检测仪上绘出稳定的基线后再停止，以便下次使用。

③ 测定鸡卵类黏蛋白的活性时，最好选用双光束紫外分光光度计。一束光路上放置空白参比样品，另一束光路则放置加有待测鸡卵类黏蛋白活性的样品，这样就可以很容易地测得 ΔA_{253nm}/min。

三、蛋白质的定量

蛋白质的定量分析是生物化学和其他生命学科最常涉及的分析内容。在生化实验中，对样品中的蛋白质进行准确可靠的定量分析，是经常进行的一项非常重要的工作。蛋白质是一种十分重要的生物大分子，它的种类很多，结构不均一，分子量又相差很大，功能各异，这样就给建立一个理想而又通用的蛋白质定量分析的方法带来了许多具体的困难。目前测定蛋白质含量的方法有很多种，下面列出根据蛋白质不同性质建立的一些蛋白质测定方法。

① 根据物理性质：紫外分光光度法。

② 根据化学性质：凯氏定氮法、双缩脲法、Folin-酚试剂法（Lowry 法)、BCA 法、胶体金法。

③ 根据染色性质：考马斯亮蓝染色法、银染法。

④ 根据其他性质：荧光法。

蛋白质测定的方法很多，但每种方法都有其特点和局限性（表 2-4)，因而需要在了解各种方法的基础上根据不同情况选用适当的方法，以满足不同的要求。

表 2-4 常用的测定蛋白质含量方法的比较

方法	测定范围 /μg・ml^{-1}	不同种类蛋白的差异	最大吸收波长/nm	特点
凯氏定氮法		小		标准方法，准确，操作麻烦，费时，灵敏度低，适用于标准的测定
紫外分光光度法	50～1000	大	280	灵敏，快速，不消耗样品，核酸类物质有影响
双缩脲法	1000～10000	小	540	重复性、线性关系好，灵敏度低，测定范围窄，样品需要量大
Folin-酚试剂法	20～500	大	750	灵敏，费时较长，干扰物质多
考马斯亮蓝 G-250 染色法	50～500	大	595	灵敏度高，简单，误差较大，颜色会转移
BCA 法	50～1000	大	562	灵敏度高，稳定，干扰因素少，费时较长
胶体金法	1～10	大	595	灵敏度极高，稳定，操作简单，干扰物影响较大

1. Folin-酚试剂法（又名 Lowry 法）

目前实验室较多采用 Folin-酚试剂法测定蛋白质含量，此法的原理是蛋白质首先在碱性

溶液中形成铜-蛋白质复合物，然后这一复合物可以与酚试剂中的磷钼酸、磷钨酸作用，产生深蓝色的钼蓝和钨蓝复合物，这种深蓝色的复合物在745～750nm处有最大的吸收峰，颜色的深浅与蛋白含量成正比。其特点是灵敏度高，较双缩脲法高两个数量级，较紫外法略高，操作稍微麻烦，反应约在15min有最大显色，并最少可稳定几个小时。其不足之处是干扰因素较多，有较多种类的物质都会影响测定结果的准确性。

2. 考马斯亮蓝G-250染色法

此方法是1976年Bradford建立，故又称为Bradford检测法。染料结合法测定蛋白质的优点是灵敏度较高，可检测到微克级的微量蛋白，操作简便、快速，试剂配制极简单，重复性好，但干扰因素较多。

考马斯亮蓝G-250具有红色和青色两种色调，在酸性溶液中游离状态下为棕红色，当它通过疏水作用与蛋白质结合后，变成蓝色，最大吸收波长从465nm转移到595nm处；在一定的范围内，蛋白质含量与595nm的吸光度成正比，测定595nm处光密度值的增加即可进行蛋白质的定量。Bio-Rad公司的蛋白质定量检测试剂盒即以此法为依据。

3. 紫外分光光度法

紫外光谱吸收法测定蛋白质含量是将蛋白质溶液直接在紫外分光光度计中测定的方法，不需要任何试剂，操作很简便，而且样品可以回收。

蛋白质溶液在280nm附近有强烈的吸收，这是由于蛋白质中酪氨酸、色氨酸残基而引起的，所以光密度受这两种氨基酸含量的支配。另外，核蛋白或提取过程中杂有的核酸对测定结果引进极大误差，其最大吸收在260nm。所以，同时测定280nm及260nm两种波长的吸光度，通过计算可得较为准确的蛋白质含量。

该方法操作简便，样品溶液可回收，同时可估计核酸含量。但核酸含量大于20%或溶液浑浊，则测定结果误差较大。在计算时也应注意各种蛋白质和各种核酸在280nm及260nm处的光吸收值也不尽相同，故计算结果有一定误差。

4. 胶体金法

采用胶体金来定量测定微量蛋白质的含量，是一种较为合适的方法。在灵敏度方面它可以达到毫微克的水平，在操作上简便易行。

胶体金是一种带负电荷的疏水性胶体。加入蛋白质后，鲜红色的胶体金溶液转呈为蓝色，最大吸收波长由490nm转移到595nm，由于颜色的改变与加入的蛋白质有定量关系，这种波长右移的现象可用于蛋白质含量的测定。

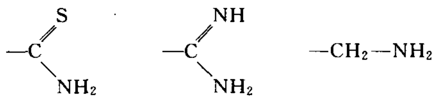
胶体金法测定蛋白质含量的优点是灵敏度高，比Lowry法、染色法等都灵敏，达到毫微克水平，胶体金试剂制备简便，操作步骤简单，反应物稳定。其缺点与其他方法相似，例如干扰物的影响和各种蛋白质的反应度不同等。

5. 双缩脲法

双缩脲（$H_2NOC—NH—CONH_2$）在碱性溶液中能与铜离子作用，形成紫红色的络合物，此反应称为双缩脲反应，蛋白质分子中含有许多肽键，与双缩脲结构类似，亦能与双缩脲试剂作用，产生颜色反应，颜色深浅与蛋白质浓度成正比。

该法灵敏度较低，一般可测定样品的蛋白质浓度范围在1～10mg/ml；必须在显色后30min内进行比色测定；30min后可能有雾状沉淀发生。

双缩脲反应并非蛋白质所特有，因此有许多干扰因素影响测定结果。凡分子内有两个或两个以上肽键的化合物以及分子内有以下结构的化合物，双缩脲反应也呈阳性。



6. BCA 法

BCA（bicinchoninic acid，二喹啉甲酸）是近来新研制的一种用于蛋白质含量检测的试剂。该方法是一种改进的 Lowry 法，反应简单，而且几乎没有干扰物质的影响。

其原理是在碱性环境下蛋白质分子中的肽键结构能与 Cu^{2+} 生成络合物，同时将 Cu^{2+} 还原成 Cu^{+}。而 BCA 试剂可敏感特异地与 Cu^{+} 结合，形成稳定的有颜色的复合物，在 562nm 处有最大吸收值。颜色的深浅与蛋白质的浓度成正比。该法的缺点是反应时间长。

第二节　DNA 的提取与分离

一、质粒 DNA 的提取与分离

质粒是携带外源基因进入细菌中扩增或表达的重要载体，这种载体在基因工程中具有非常广泛的应用价值，而质粒的分离与提取则是分子生物学最常用、最基本的实验技术。已经提出过许多方法用于从细菌中提纯质粒 DNA，这些方法都含有以下 3 个步骤：细菌的培养，细菌的收获和裂解，质粒 DNA 的分离与纯化。

（一）质粒提取的一般步骤

1. 细菌的培养

将分离的一个单菌落，接种到培养物中（含有适当抗生素的液体培养基）生长，随着细菌的生长，质粒 DNA 也在自主复制，然后从中纯化质粒。现在使用的许多质粒载体（如 pUC 系列）都能复制到很高的拷贝数，只要培养物在标准 LB 培养基中生长到对数晚期，就可以大量提纯质粒。然而，较早一代的载体（如 pBR322）由于不能如此自由地复制，所以需要在得到部分生长的细菌培养物中加入氯霉素继续培养若干小时，以便对质粒进行扩增。氯霉素可抑制宿主的蛋白质合成，结果阻止了细菌染色体的复制；然而，松弛型质粒仍可继续复制，在若干小时内，其拷贝数持续递增。这样，像 pBR322 一类的质粒，从经氯霉素处理和未经处理的培养物中提取质粒的产量迥然不同，前者大为增高。多年来，加入足以完全抑制蛋白质合成的氯霉素已成为标准的操作。该方法提取的质粒 DNA 量可满足分子克隆中几乎所有的工作任务。

2. 细菌的收获和裂解

细菌的收获可通过离心来进行，而细菌的裂解则可以采用多种方法中的任意一种，这些方法包括用非离子型或离子型去污剂、有机溶剂或碱进行处理及用加热处理等。选择哪一种方法取决于 3 个因素：质粒的大小、大肠杆菌菌株及裂解后用于纯化质粒 DNA 的技术。

虽然不可能针对质粒和宿主的每一种组合分别提出精确的裂解条件，但仍可据下述一般准则来选择适当方法，以取得满意的结果。

（1）质粒的大小　大于 15kb 的质粒容易受损，故应采用温和裂解法从细胞中释放出

来。将细菌悬于含有蔗糖的等渗溶液中，以减轻细菌裂解时 DNA 释出过快产生的机械剪切力。然后用溶菌酶和 EDTA 进行处理，破坏细胞壁和细胞外膜，再加入 SDS 一类去污剂，解聚蛋白质与 DNA 的结合。这种方法最大限度地减小了操作中的物理剪切力。分离小质粒可采用较剧烈的方法。在加入 EDTA 后，或在加入溶菌酶后让细菌暴露于去污剂中，通过煮沸或碱处理使之裂解。这些处理可破坏碱基配对，故可使宿主的线状染色体 DNA 变性，但闭环质粒 DNA 链由于处于拓扑缠绕状态而不能彼此分开。当条件恢复正常时，质粒 DNA 链迅速得到准确配对，重新形成完全天然的超螺旋分子。

（2）菌株的类型　一些大肠杆菌菌株（如 HB101 的一些变种衍生株）富含糖类，用去污剂或加热裂解时糖类可释放出来，若用氯化铯-溴化乙锭梯度平衡离心进行质粒纯化时，糖类会在梯度中紧靠超螺旋质粒 DNA 所占位置形成一致密的、模糊的区带。因此很难避免质粒 DNA 内污染有糖类，而糖类可抑制多种限制酶的活性。因此从诸如 HB101 和 TG1 等大肠杆菌菌株中大量制备质粒时，不宜使用煮沸法。当从表达内切核酸酶 A 的大肠杆菌菌株（$endA^{+}$ 株，如 HB101）中小量制备质粒时，建议也不使用煮沸法。因为煮沸不能完全灭活内切核酸酶 A，以后在温育（如用限制酶消化）时．质粒 DNA 会被降解。若一定要用煮沸法提取这类菌株的质粒 DNA，必须用酚、氯仿反复抽提几次，以去除该酶。

（3）氯霉素的应用　新一代质粒的拷贝数都非常高，以致于不需要用氯霉素进行选择性扩增就可获得高产。然而，某些工作者沿用氯霉素并不是要增加质粒 DNA 的产量，而是要降低细菌细胞在用于大量制备的溶液中所占体积。大量高度黏稠的浓缩细菌裂解物，处理起来煞为费事，而在对数中期在培养物中加入氯霉素可以避免这种现象。有氯霉素存在时从较少量细胞获得的质粒 DNA 的量与不加氯霉素时从较大量细胞所得到的质粒 DNA 的量大致相等。

3. 质粒 DNA 的纯化

质粒从细菌中分离出来以后，多数时候需要进一步纯化。常使用的所有纯化方法都利用了质粒 DNA 相对较小及共价闭合环状这样两个性质。例如，用氯化铯-溴化乙锭梯度平衡离心分离质粒和染色体 DNA 就取决于溴化乙锭与线状 DNA 分子和闭环 DNA 分子的结合量有所不同。多年来，氯化铯-溴化乙锭梯度平衡离心是制备大量质粒 DNA 的可靠经典方法。然而该过程既昂贵又费时，为此发展了许多替代方法，其中主要包括利用离子交换层析、凝胶过滤层析、分级沉淀等分离质粒 DNA 和宿主 DNA 的方法，也可获得较好质量的质粒。

（二）质粒 DNA 的小量制备

质粒 DNA 的小量制备通常采用碱变性法或煮沸法。

1. 碱变性裂解法快速提取小量质粒 DNA

碱变性法抽提质粒是基于染色体 DNA 与质粒 DNA 的变性与复性的差异而达到分离目的。在 pH12.6 的高碱性条件下，染色体 DNA 和质粒 DNA 都发生变性，但质粒超螺旋共价闭合环状结构的两条互补链不会完全分离，所以当以 pH4.8 的乙酸钾高盐缓冲液调节其 pH 至中性时，变性的质粒 DNA 又恢复到原来的构型，保存在溶液中；而染色体 DNA 不能复性而形成缠连的网状结构，通过离心，染色体 DNA 与不稳定的大分子 RNA、蛋白质-SDS 复合物等一起沉淀下来而被除去。

2. 煮沸法快速提取小量质粒 DNA

大肠杆菌细胞经溶菌酶和 Triton X-100 处理后，在高温条件下使蛋白质变性。变性蛋

白带着染色体 DNA 一起沉淀下来，质粒 DNA 仍留在上清液中。用异丙醇或乙醇处理可以沉淀 DNA，获得质粒 DNA 的制品。

（三）质粒 DNA 的大量制备

对转化成功的重组质粒进行分析鉴定后，需对重组质粒进行大量扩增制备，以获得足够量的基因载体 DNA 用于基因转移实验。有多种方法可用于大量制备质粒 DNA。这些方法各有所长，但均包括三个基本步骤：细菌的培养，细菌的收集和裂解，质粒 DNA 的分离和纯化。

从琼脂平板上挑取一个单菌落，接种到含有适当抗生素的液体培养基中培养，然后通过离心回收菌体，即可从中纯化质粒。对于拷贝数较低的质粒（如 pBR322 等），有必要使用氯霉素（170μg/ml）抑制宿主蛋白质的合成，使细菌细胞内质粒的拷贝数得以较大幅度地增加。现在使用的许多质粒载体如 pUC 系列，由于具有很高的拷贝数，无需选择性地扩增质粒 DNA。但在对数中期向培养物中加入氯霉素，可抑制宿主细胞蛋白质的合成，使宿主细胞的代谢限制在一定水平，从而大幅度地降低细菌裂解物的黏度，因此，在保证获得与不加氯霉素等量的质粒 DNA 的同时，使随后的提取和纯化工作简单易行。此外，对于携带某些较大质粒而生长较差的细菌，可用丰富培养基（如 Terrific 肉汤）改善其生长，以获得更高的质粒 DNA 产量（通常可使产量增加 3～5 倍）。

细菌细胞裂解的方法很多，但共同的步骤是用溶菌酶和 EDTA（乙二胺四乙酸）处理，破坏细胞壁和细胞膜，再加入 SDS（十二烷基硫酸钠）一类的去污剂溶解球形体，使之发生温和的溶菌作用。于是，质粒 DNA 分子、多聚核糖体及其他可溶性物质从细胞中逐步释放出来。而染色体 DNA 大分子与细胞碎片缠绕在一起，可经离心与质粒 DNA 分开。如果待分离的质粒较小，还可以用更剧烈的方法裂解细胞，如按上述程序加入去污剂后，通过煮沸或碱处理可更有效地裂解细胞。这些处理可能会破坏 DNA 分子的碱基配对，故可使宿主的线状染色体 DNA 分子变性，但闭环质粒 DNA 链由于处于拓扑缠绕状态而不能彼此分开。当条件恢复正常时，质粒 DNA 链迅速得到准确的配对，重新形成完全天然的超螺旋分子。对于较大的质粒（大于 15kb），由于其容易受损，故应采取温和的裂解方法使之从细胞中释放出来。为此，可将细菌悬浮于蔗糖等渗溶液中，然后再用溶菌酶和 EDTA 进行处理，破坏细菌的细胞壁和外膜，再加入 SDS 一类的去污剂溶解球形体。这种方法最大限度地减小了从具有正压的细菌内部把质粒释放出来时所需的作用力。

将离心所得的含有质粒 DNA 的上清液，经有机溶剂（如异丙醇、乙醇等）沉淀可回收到质粒 DNA 粗品，其中含有大量的蛋白质和 RNA 等，需经进一步纯化。目前采用的纯化方法有三种：氯化铯密度梯度离心法、Sepharose 2B 柱色谱法和氯化锂-聚乙二醇沉淀法。聚乙二醇沉淀法操作简便，可有效地纯化碱裂解法获得的质粒 DNA，但缺点是不能有效地将带有缺口的环状分子同闭环质粒 DNA 分开。因此，尽管氯化铯密度梯度离心法既昂贵又费时，但在某些情况下仍是首选方法。

1. 质粒的抽提（碱变性法）的操作过程

① 将大肠杆菌的一个单菌落接种入盛有 200ml LB 培养基（含 100μg/ml 氨苄青霉素）的 500ml 三角瓶中，置 37℃振荡培养过夜（置摇床中，150r/min）。

② 将培养物转入 50ml 塑料离心管内，于室温 3000r/min 离心 10min。

③ 弃上清，将离心管倒立，使上清流净，用纸巾或吸水纸将液体吸干。

④ 重复步骤②～③，每管中收获有 100ml 培养液的细菌，加 5ml GET 溶液［终浓度

为：500mol/L 葡萄糖、10mol/L EDTA、25mol/L Tris-HCl（pH8.0），高压灭菌，4mg/ml 溶菌酶（临用前加入）]。悬浮细菌后转入 50ml 塑料离心管（1 管）。2500r/min 于室温离心 10min。

⑤ 弃上清。加入 2ml 溶液Ⅰ［50mmol/L Tris-HCl（pH8.0），10mol/L EDTA］，悬浮细菌。

⑥ 加入 4ml 溶液Ⅱ（0.2mol/L NaOH，1% SDS），盖上盖子，将离心管颠倒（约 7 次），混匀管内液体，边颠倒边旋转离心管，不要用旋涡振荡器，将离心管置冰浴 5min。

⑦ 每管加入 3ml 冰中预冷的溶液Ⅲ（5mol/L 乙酸钾：冰乙酸：水，按 6：1.15：2.85 沸合而成，所配溶液对钾是 3mol/L，对乙酸根是 5mol/L），振荡离心管数次，使溶液Ⅲ充分分散到黏稠的细菌裂解物中，将离心管置冰浴 10min。

⑧ 12000r/min 离心 10min，将上清转入另一支 50ml 离心管，加入 0.6 体积的异丙醇，混匀，室温放置 5～10min。

⑨ 12000r/min 离心 10min，弃上清。加入 1ml TE 缓冲液［10mmol/L Tric-HCl（pH7.4），1mol/L EDTA］，将沉淀溶解，此为质粒 DNA 粗提液。

2. 质粒的纯化

（1）氯化锂-聚乙二醇沉淀法

① 取 1ml 质粒 DNA 粗提液，加入等体积的 5mol/L LiCl，混匀；置冰浴 10min。

② 12000r/min 离心 10min，将上清转入另一支离心管中，加入等体积的异丙醇。置室温 10min。

③ 12000r/min 离心 10min，弃上清，将沉淀溶于总体积为 200μl 的 TE 缓冲液中，转入一个 Eppendorf 管。

④ 可选择：重复步骤①～②一次。最终仍将沉淀溶于 200μl TE 缓冲液中。

⑤ 加入 5μl RNase A 溶液（2μg/μl，用 TE 缓冲液稀释 10mg/ml 贮存液而成），混匀，置 37℃温育 30min。

⑥ 加入等体积（200μl）的 1.6mol/L NaCl-13% PEG，混匀后置冰浴 10～20min。

⑦ 12000r/min，于室温离心 5min。

⑧ 吸弃上清，将沉淀溶于 200μl TE 缓冲液中，用酚-氯仿-异戊醇（25：24：1，体积比）抽提 2 次，再用氯仿-异戊醇（25：1，体积比）抽提 1 次。

⑨ 将水相转入一个新的 Eppendorf 管，加入 0.1 倍体积的 3mol/L 乙酸钠（pH 5.2）和 2 倍体积－20℃预冷的无水乙醇。置－20℃ 10～20min。

⑩ 12000r/min，室温离心 5min。弃上清。将 DNA 溶于 100μl TE 缓冲液中，置 4℃或－20℃贮存。

（2）CsCl 密度梯度超离心法

① 在 5ml 质粒粗提液中加入 3ml TE 缓冲液，使总体积为 8ml。

② 用 2mol/L Tris 溶液调节 pH 至 7.0。

③ 加入 9g CsCl。

④ 加入 0.8ml 溴化乙锭（10mg/ml）。

⑤ 在 4℃，5000r/min，离心 10min。

⑥ 将上清转入 VTi 50 的超离心管中，沉淀弃去。

⑦ 用 TE 缓冲液补满离心管，仔细平衡后，将离心管严密封口。

⑧ 在超速离心机中于25℃，40000r/min离心42h。离心后，在一般光线下应能见到2条DNA带。上面的区带含有线性DNA和有缺口的环状质粒DNA，下面的区带则含有闭合环状质粒DNA。

⑨ 在紫外灯下，用注射器抽出下面的区带（此为闭合环状DNA），转入10ml试管中。

⑩ 加入3倍体积的双蒸水稀释。

⑪ 加入2倍体积的冷冻无水乙醇，在－20℃置放至少2h以上。

⑫ 于0℃，10000r/min离心30min，弃去上清。

⑬ 用5ml 70%乙醇洗沉淀1次，同上再离心，去除上清。

⑭ 将沉淀溶解在500μl TE（pH8.0）中，将溶液转入一Eppendorf管，加入2.5μl RNase A（10mg/ml），在37℃保温30min。

⑮ 用等体积苯酚-氯仿（1∶1，体积比）抽提1次，再用等体积氯仿抽提1次，每次将水相转入一新的Eppendorf管。

⑯ 加入1/10体积3mol/L乙酸钠（pH 4.8）和2倍体积的无水乙醇，混匀，静置10min后10000r/min，4℃离心5min。

⑰ 倾除上清，用70%乙醇清洗沉淀，10000r/min，4℃离心5min。

⑱ 将沉淀溶于250μl TE缓冲液中，取少量样品进行DNA含量测定。

二、染色体DNA的提取与分离

一般地说，总DNA主要是指基因组DNA(genomic DNA)，即细胞核内的染色体DNA分子。核DNA分子呈极不对称的线状结构，一条染色体为一个DNA分子，为细长的纤丝状分子，使其对机械力十分敏感。虽然目前可以成功地测定出它的一级结构，但仍难以分离出它的完整分子。分离纯化过程中，DNA分子的断裂是很难避免的。在进行基因组文库的构建及Southern印迹实验时，应尽量分离出片段较长的DNA分子。

真核细胞的破碎有多种手段。物理方式包括超声波法、匀浆法、液氮破碎法等，这些物理方法容易导致DNA链的断裂。为了尽可能获得高分子量的DNA，一般采用去污剂温和处理法，对于细胞破碎较困难的材料，可以辅加蛋白酶K，在酶的存在下共同破碎细胞。下面介绍几种哺乳动物细胞高分子量DNA的提取方法。

（一）从培养细胞中提取DNA

① 吸取培养液，用4℃预冷的Tris缓冲盐溶液洗细胞1次。

② 胰酶消化收集细胞，或用橡皮刮把细胞从培养瓶壁上刮下。将细胞悬浮在少量（约0.5ml）4℃预冷的TE中。

③ 加10倍体积的细胞破碎液［0.5mol/L的EDTA（pH8.0），100μg/ml蛋白酶K，0.5% SDS］，混匀后，37℃过夜，或50℃水浴保温3h。

④ 用等体积的酚抽提1次。2500r/min离心，收集水相，用1/2体积的TE缓冲液饱和酚加1/2体积的氯仿-异戊醇（24∶1，体积比）抽提1次。2500r/min离心，收集水相。等体积氯仿-异戊醇（25∶1，体积比）抽提1次。

⑤ 收集水相，装入透析袋（透析袋处理：先放在2%碳酸氢钠-1mmol/L EDTA溶液中煮沸10min，重蒸水冲洗后再在重蒸水中煮沸10min。存放在50%乙醇或重蒸水中，4℃保存。使用前用透析液冲洗透析袋内壁）。放入透析液［50mmol/L Tris-HCl（pH8.0），10mmol/L EDTA，10mmol/L NaCl］。4℃过夜。

⑥ 吸出 DNA 溶液，加入 RNase A 到终浓度为 100μg/ml，置 37℃温育 3h。

⑦ 加 SDS 至终浓度为 0.5%，然后加蛋白酶 K 至 100μg/ml，37℃过夜，或 50℃ 3h。

⑧ 酚-氯仿（1∶1，体积比）、氯仿-异戊醇（24∶1，体积比）各抽提 1 次。

⑨ 透析，4℃过夜。换 1、2 次透析液［10mmol/L Tris-HCl（pH8.0），1mmol/L EDTA］。

⑩ 用分光光度计测定透析液的 A_{279}，读数要小于 0.05，否则再换新鲜透析液进行透析，以除去酚等杂质。

⑪ 测定 DNA 溶液的 A_{260} 和 A_{280} 值，A_{260}/A_{280} 应大于 1.8，比值偏低表明制备物中蛋白质含量偏高。在这种情况下，重复实验步骤⑧～⑪。

⑫ 计算 DNA 浓度（$A_{260}=1$ 时，溶液所含 DNA 约为 50μg/ml），通过电泳分析一小份样品，以鉴定 DNA 有无降解，初步测定 DNA 的分子量。

（二）从组织中提取 DNA

① 把组织块剪碎，放在液氮中。

② 把冻成硬粒的组织块放入不锈钢制的捣碎器，加液氮，把组织块捣碎成粉末状。

③ 待液氮挥发后，加入少量预冷的 TE，要防止组织碎块聚结成团，尽可能地制成细胞悬液。

④ 如果是肝脏等组织，可直接用匀浆器（管）制成匀浆，悬浮在 TE 中，匀浆时应注意将匀浆器置于冰内。

⑤ 按照上述“（一）从培养细胞中提取 DNA”方法的步骤③～⑪提取 DNA。

（三）从大量血样的白细胞中分离高分子量 DNA

① 抽 20ml 血，肝素抗凝。

② 加 20ml 1.8% NaCl，加 20ml 6%葡聚糖（Dextran），于室温直立静置。

③ 取上清液，离心，1000r/min，4℃，10min。

④ 取沉淀细胞，悬浮打散后加入 30ml 重蒸水，再加 10ml 3.6% NaCl。

⑤ 离心，1000r/min，4℃，10min。

⑥ 将沉淀的细胞悬浮在 10ml 0.5mol/L EDTA 中，加入 20% SDS 至终浓度 1%，加入 10mg/ml 蛋白酶 K 至终浓度 100μg/ml。

⑦ 轻轻摇动，37℃，过夜。

⑧ 用酚，酚-氯仿-异戊醇（25∶24∶1，体积比），氯仿-异戊醇（24∶1，体积比）各抽提 1 次。

⑨ 加 5mol/L NaCl 至终浓度为 0.1mol/L，混匀。

⑩ 加 2 倍体积－20℃预冷的无水乙醇，混匀。置－20℃或－70℃，过夜。

⑪ 离心，12000r/min，4℃，15min；倾去上清，待乙醇挥发后，将 DNA 沉淀溶于 TE（5～10ml）。

⑫ 待 DNA 完全溶解后，加 RNase A 至 100μg/ml，置 37℃温育 3h。

⑬ 加 SDS 至 0.5%，再加蛋白酶 K 至 100μg/ml，37℃，过夜。

⑭ 用酚，酚-氯仿-异戊醇（25∶24∶1，体积比），氯仿-异戊醇（24∶1，体积比）各抽提 1 次。

⑮ 取上层水相，装入透析袋，透析液为 10mmol/L Tris-HCl（pH8）、1mmol/L EDTA。于 4℃透析过夜，换 1～2 次透析液。

⑯ 测 DNA 浓度。

（四）从小量血样的白细胞中分离高分子量 DNA

① 取 2ml 血，用 EDTA 或柠檬酸抗凝。

② 加入 6ml 细胞裂解液（155mmol/L NH_4Cl，10mmol/L $KHCO_3$，0.1mmol/L EDTA），置于冰浴中 10min，使红细胞裂解。

③ 离心，1000r/min，4℃，10min。白细胞沉淀重新悬浮于 2ml 0.15mol/L NaCl-0.015mol/L 柠檬酸钠-20mmol/L EDTA 中。

④ 加入蛋白酶 K 200μg，再加入 10% SDS 至终浓度为 1%，37℃保温 2h。

⑤ 用等体积的酚-氯仿-异戊醇（25∶24∶1，体积比）抽提 2 次，氯仿-异戊醇（24∶1，体积比）抽提 1 次。

⑥ 加入 1/30 体积的 3mol/L 乙酸钠（pH5.2）和等体积的异丙醇，混匀，10000r/min 离心 15min。沉淀用 70%乙醇洗 1 次，沉淀溶于 3ml TE 缓冲液中。

⑦ 加入 RNase A 至终浓度 20μg/ml，37℃保温 2h。加入蛋白酶 K 至终浓度 100μg/ml，再加入 10% SDS 使终浓度为 0.5%，37℃保温 2h。

⑧ 加入等体积的酚-氯仿-异戊醇（25∶24∶1，体积比）抽提 2 次，再用氯仿-异戊醇（24∶1，体积比）抽提 1 次。

⑨ 取上层水相，加入等体积的异丙醇，混匀。12000r/min 离心 10min。沉淀溶于 0.5ml TE 缓冲液中，置－20℃冰箱中备用。

（五）高分子量 DNA 的小量快速制备

① 将细胞悬液浓度调整成 1×10^6 个细胞/ml，取 0.5ml 转入 Eppendorf 管内，3000r/min 离心 5min，弃上清。

② 将细胞悬于 0.5ml TES [150mmol/L NaCl，10mmol/L Tris-HCl（pH7.5），1mmol/L EDTA] 中，加入 10% SDS 0.1ml，置室温 5min。

③ 加入等体积酚-氯仿-异戊醇（25∶24∶1，体积比）抽提 1 次，再用氯仿-异戊醇（24∶1，体积比）抽提 1 次。

④ 将上层水相转入另一 Eppendorf 管，加入 1/10 体积的 3mol/L 乙酸钠（pH5.2），2 倍体积的无水乙醇，混匀后即可见团状 DNA 沉淀。

⑤ 10000r/min 离心 1min，弃上清。

⑥ 将 DNA 溶于 0.1ml TE 缓冲液，置－20℃保存。该 DNA 制品可用于 Southern 印迹或用作 PCR 反应的模板。

第三节 RNA 的提取与分离

RNA 主要由以下几类分子组成：rRNA（ribosomal RNA，核糖体 RNA），占总量的 80%～85%；tRNA（transfer RNA，转移 RNA）和核内小分子 RNA 两者占 10%～15%；mRNA（messerger RNA，信使 mRNA）占 1%～5%。rRNA 在总 RNA 分子中含量丰富。

真核细胞的 rRNA 由 28S、18S、5S 及 4S 几类组成［S 为沉降系数（*s*）的单位，*s* 的定义是离心场强度的沉降强度，用于近似地表示生物大分子的大小］。mRNA 分子种类繁多，分子量大小不一（长度为数百个至数千个碱基），大多数真核生物的 mRNA 的 3′末端都带有一段 20～250 个多聚腺苷酸（polyA）。利用此特征可以很方便地从总 RNA 中用寡聚（dT）亲和色谱的方法分离 mRNA。

从细胞中分离的 RNA 的纯度与完整性对于许多分子生物学实验至关重要，如 Northern 印迹杂交分析、寡聚（dT）纤维素选择分离 mRNA、cDNA（complementary DNA，互补 DNA）合成、体外翻译及 mRNA 差别显示技术分离基因等实验，在很大程度上取决于 RNA 的质量。

RNA 分离的关键因素是尽量减少 RNA 酶的污染，因此，在 RNA 制备过程中，必须注意抑制 RNA 酶的活性。由于 RNA 很容易被 RNA 酶降解，加上 RNA 酶极为稳定而且广泛存在，因此在提取过程中要严格防止 RNA 酶的污染。

为了避免 RNA 酶污染，应严格采取以下措施。①溶液处理。RNA 制备过程中使用的所有溶液（包括水和其他盐溶液）都必须先用焦碳酸二乙酸脂（DEPC）进行处理，DEPC 可通过共价结合灭活 RNA 酶。但要注意含 Tris 碱的溶液用 DEPC 处理效果不好，因为 Tris 碱可与 DEPC 发生反应并使之失活。②玻璃和塑料器皿的处理。RNA 制备过程中使用的实验器皿应先进行处理以除去残存的 RNA 酶活性。高压处理不能完全灭活 RNA 酶，因此，玻璃器皿应在 300℃下干烤。塑料器皿则应用氯仿冲洗，可达到灭活 RNA 酶的目的。如果是从完整包装中新取出的塑料制品，一般无 RNA 酶污染，可以直接使用。③实验者的双手是 RNA 酶的重要污染源，因此务必戴手套进行操作。

一、组织培养细胞胞质 RNA 的制备

实验中首先用预冷的磷酸盐缓冲液（PBS）洗涤细胞，将收集的细胞重悬于裂解缓冲液（含非离子型去污剂 NP-40）中，细胞质膜迅速裂解，再通过短暂的离心去除完整的细胞核。回收上清胞质液，并加入 SDS 使蛋白质变性，加入蛋白酶消化后，再用酚-氯仿和氯仿 2 次抽提以除去蛋白质。最后用乙醇沉淀得到胞质 RNA，通过测定 A_{260} 和 A_{280} 进行定量。注意所有操作始终在冰上（或冰浴中）进行。

1. 基本方案操作步骤

① 用预冷的 PBS 缓冲液（137mmol/L NaCl，2.7mmol/L KCl，4.3mmol/L $Na_2HPO_4 \cdot 7H_2O$，1.4mmol/L KH_2PO_4）洗涤细胞。

单层培养细胞，经预冷的 PBS 洗涤 3 次后，在直径 100mm 培养皿中加入 1ml 预冷的 PBS（直径 150mm 培养皿加入 3～5ml PBS），用细胞刮子将细胞刮落，转移至预冷离心管，300×*g* 离心 5min，收集细胞，置于冰上。悬浮培养细胞 300×*g* 5min 离心沉淀后，重悬于原培养液一半体积的预冷 PBS 中，再次离心沉淀。

② 重悬细胞于 375μl 裂解缓冲液［50mmol/L Tris-HCl（pH8.0），100mmol/L NaCl，5mmol/L $MgCl_2$，0.5%（体积分数）NP-40，用 DEPC 处理的水配制，过滤灭菌。如制备的 RNA 是用于 Northern 印迹分析或细胞富含 RNA 酶，则在裂解缓冲液中加入 RNA 酶抑制剂：100U/ml 胎盘核糖核酸酶抑制剂（如 RNasin）、1mmol/L DTT 或 10mmol/L 氧钒核苷复合物］中，冰浴 5min。悬液应迅速变清亮，提示细胞裂解充分。此操作一方面要剧烈振摇以使细胞裂解完全，一方面又要小心避免出现气泡。

③ 将细胞转移至微量离心管中，4℃离心 2min。上清即为细胞胞质提取物，一般略呈云雾状或为黄白色（因细胞不同而不同）；沉淀为细胞核和一些细胞碎片，为白色物，量应少于步骤①中总细胞沉淀量。

④ 转移上清至含 4μl 20% SDS 的干净试管中，立即振荡混匀。

⑤ 加入 2.5μl 蛋白酶 K（20mg/ml），37℃温育 15min。

⑥ 加入 400μl 酚-氯仿-异戊醇（25∶24∶1，体积比）抽提，充分振荡 1min 以上，10000r/min 离心 5min（或以上）。提取和离心过程在常温下完成。

注：经蛋白酶处理后再抽提，两相间界面的沉淀量应很少，当然这与细胞类型也有关。对某些细胞可以省略蛋白酶处理一步，此时界面间的白色沉淀量可能很多，这样除去有机相后，可回收的水相就很少。如出现这种情况，首先可尝试再离心数分钟；若白色沉淀物仍不能压缩至界面，就从管底吸除有机相，再加入 400μl 氯仿-异戊醇（24∶1，体积比），充分振荡后离心 2min，经此处理可使沉淀物大部分消失。回收水相再进行以下步骤。

⑦ 回收上层水相并转移到一干净试管中，注意不要吸到界面的沉淀物。加入 400μl 酚-氯仿-异戊醇（25∶24∶1，体积比），重复抽提一次。

⑧ 同样回收上层水相并转移到一干净试管中，加入 400μl 氯仿-异戊醇（24∶1，体积比），振荡 15～30s，10000r/min 离心 1min。再回收上层水相，转移到一干净试管中。

⑨ 加入 4μl 3mol/L 乙酸钠（pH5.2）（即 1/10 体积，终浓度为 0.3mol/L）和 1ml 无水乙醇，上下颠倒混匀，冰浴 15～30min 或放在－20℃过夜。

⑩ 4℃ 10000r/min 离心、15min，回收 RNA。

⑪ 用 1ml 75%乙醇-25% 0.1mol/L 乙酸钠（pH5.2）洗涤 RNA 沉淀。

⑫ 干燥后用 100μl DEPC 处理过的水重悬 RNA。测定浓度时取 10μl 稀释于 990μl 水中，测 A_{260} 和 A_{280} 值，其余样品－70℃保存备用。

本操作适用于 2×10^7 个细胞（约 2 盘直径 10cm 培养皿细胞，或 20ml 悬浮培养细胞）。如细胞数量更多，可按比例增加各溶液用量。

注：DEPC 处理溶液时，每 100ml 待处理溶液加入 0.2ml DEPC，剧烈振荡混匀，再经高压灭活 DEPC，单独存放，用于 RNA 实验中的溶液配制。DEPC 为可疑致癌物，应戴手套并在通风橱中操作。

本方案适用于大多数细胞类型，但没有针对核酸酶采取特殊措施，因此对于某些细胞类型，据此方案分离的 RNA 质量可能不能用于 Northern 杂交。如果要得到全长 RNA，裂解缓冲液中应加入 RNA 酶抑制剂或采用胍盐法。此外，如果 RNA 是从瞬时转染 DNA 的细胞中进行分离，则应用 DNA 酶处理以除去转染的 DNA。如果 RNA 要用于采用均匀标记的探针进行核酸酶保护实验分析，此过程尤为重要。

操作应迅速，所有步骤均在冰上完成，直至在胞质提取物中加入 SDS，这样可以很好地避免 RNA 的降解。对于富含 RNA 酶的细胞，应在裂解缓冲液中加入 RNA 酶抑制剂，但在大多数情况下并不需要。对某些细胞类型可以省略蛋白酶 K 一步，而在除掉细胞核和加入 SDS 后直接进行有机溶剂抽提过程。DNA 的污染问题只有在进行瞬时表达分析、从转染 DNA 的细胞制备 RNA 时才会遇到；这种情况下应增加 DNA 酶处理步骤。

RNA 产量依据细胞不同有较大变化，一个直径 10cm 培养皿所收获的大多数纤维细胞系细胞或 10^7 个淋巴细胞可得到 30～100μg RNA，A_{260}/A_{280} 的比值应在 1.7～2.0；A_{260} 值为 1 时，RNA 含量为 40μg/ml。

2. DNA 污染的除去

用以上基本方案从瞬时转染 DNA 的细胞中分离得到的 RNA 中可能污染有大量的模板 DNA，并会干扰 RNA 的分析（如核酸酶保护实验），尤其是使用均匀标记的探针时。除掉污染 DNA，可在基本方案第⑩步后接本方案操作步骤。

试剂　100mmol/L $MgCl_2$；10mmol/L 二硫苏糖醇（DTT）；2.5mg/ml 无 RNA 酶的 DNA 酶Ⅰ；胎盘核糖核酸酶抑制剂［如 Promega 公司的 RNasin 或氧钒核苷（Vandyl-ribonucleoside）复合物］；DNA 酶反应终止液（50mmol/L EDTA，1.5mol/L 乙酸钠，1% SDS）。

① RNA 重新溶解于 50μl TE 缓冲液［10mmol/L Tris-HCl（pH7.4），1mmol/L EDTA］中。

② 在冰上配制如下混合液：10μl 100mmol/L $MgCl_2$-10mmol/L DTT，0.2μl 2.5mg/ml 无 RNA 酶的 DNA 酶Ⅰ，0.1μl 胎盘核糖核酸酶抑制剂（25～50U/μl）和 39.7μl TE 缓冲液。

③ 每份 RNA 样品加 50μl 混合液，混匀，37℃温育 15min。

④ 加 25μl DNA 酶反应终止液终止反应。

⑤ 先用酚-氯仿-异戊醇（25∶24∶1，体积比）抽提 1 次后，再用氯仿-异戊醇（24∶1，体积比）抽提 1 次。

⑥ 加入 325μl 无水乙醇，置于冰上沉淀 15～30min 或－20℃过夜。

⑦ 接基本方案操作步骤⑪。

二、胍盐法制备总 RNA

本节介绍胍盐法制备 RNA 的三种方案。第一、第二种方案采用氯化铯梯度离心法分离 RNA，第三种方案通过一步法分离 RNA，不需要进行氯化铯梯度离心。

（一）氯化铯纯化培养细胞的 RNA

培养细胞经洗涤后，加入 4mol/L 胍盐溶液裂解细胞。用 20G 针头抽吸细胞裂解物以降低溶液黏度，经氯化铯梯度离心沉淀 RNA。仔细吸除上清，将 DNA、蛋白质与沉淀中的 RNA 完全分离开。沉淀中的 RNA 溶解后，再经乙醇沉淀，最后用分光光度法测 A_{260} 进行定量。

下列①～④步均在常温下进行。

①a 对单层细胞，室温下每一培养皿用 5ml PBS 缓冲液洗涤 2 次。

①b 对悬浮培养细胞，300×g 5min 离心沉淀（细胞数≤10^8 个），除去上清，沉淀重悬于培养液一半体积的 PBS 中，再次离心沉淀，除去上清。

②a 对单层培养细胞，每 10^8 个细胞加 3.5ml 胍盐溶液［4mol/L 异硫氰酸胍，20mmol/L 乙酸钠（pH 5.2），0.1mmol/L 二硫苏糖醇（DTT），0.5%十二烷基肌氨酸钠（Sarkosyl）。在水中溶解异硫氰酸胍和乙酸钠，加热溶液（65℃）使胍盐溶解，加入 DTT 和十二烷基肌氨酸钠。冰乙酸调节溶液 pH 5.5，补足容量后用 Nalgene 滤器过滤，室温贮存溶液］，培养皿内细胞应迅速裂解，用细胞刮子将黏稠的裂解物刮到培养皿的一侧，再用带 20G 针头的注射器回收裂解物。

②b 对悬浮培养细胞，在离心管中加入 3.5ml 胍盐溶液。

③ 细胞裂解物非常黏稠，使用 20G 针头抽吸 4 次，然后回收到一干净试管中。这一步

的目的是反复剪切染色体 DNA，降低其黏度，这样离心后 DNA 可完全除去。

④ 在硅化并经高压灭菌后的同质异晶超速离心管中加 1.5ml 5.7mol/L CsCl [CsCl 溶解于 0.1mol/L EDTA（pH8.0）中，加 0.2%体积 DEPC，振摇 20～30min 后高压灭菌处理。在高压灭菌处理前后称量溶液，并在高压灭菌处理后用 DEPC 处理水补足至原重量，确保溶液终浓度为 5.7mol/L]，加入 3.5ml 的细胞裂解物至 CsCl 垫层顶部，形成阶梯梯度，界面应能看到。试管作硅化处理的目的是减少黏附在试管壁上的物质的量，从而减少对 RNA 的污染。

⑤ 用 Beckman 离心机和 SW-55 转头，以 35000r/min（150000×*g*），18℃离心 12～20h。注意缓慢加速和减速，避免破坏梯度。

⑥ 小心地除去上清液。将巴斯德吸管的底部伸进溶液顶层，边吸边随液面的下降而下降，剩下约 100μl 的时候，小心倒置试管，让残余液体流出。界面处应有一层白色物，含细胞 DNA，注意要完全除去该层。

⑦ 干燥沉淀 5～10min，重悬于 360μl TES 溶液 [10mmol/L Tris-HCl（pH7.4），5mmol/L EDTA，1%SDS] 中，用吸管反复上下吹吸，在室温下放置 5～10min，以使沉淀充分溶解，再转移到一干净离心管中。重悬沉淀的时间一定要充分，否则 RNA 的产量会大大降低。

⑧ 加 40μl 3mol/L 乙酸钠（pH5.2）和 1ml 无水乙醇，在干冰/乙醇上沉淀 30min，离心 10～15min，除去上清液。

⑨ 重悬沉淀于 360μl 水中，重复步骤⑧。

⑩ 干燥沉淀 10min，再溶解于约 200μl 水中。取 10μl 稀释至 1ml，定量，以水或乙醇溶解沉淀物，－70℃贮存。

用此方案得到的 RNA 可以进行 Northern 印迹、S1 核酸酶分析和 SP6 RNA 聚合酶分析。如果需要纯度更高的 RNA，可将步骤⑦进行以下改进：重悬沉淀于 TES 缓冲液中后，加 360μl 4∶1（体积比）氯仿-1-丁烷抽提，回收上清，剩余氯仿中再加 360μl TES 溶液抽提 1 次。合并上清，按步骤⑧加 1/10 体积 3mol/L 乙酸钠（pH5.2）和无水乙醇沉淀。

注：乙酸钠、水和 5.7mol/L CsCl 都应用 DEPC 处理以灭活 RNA 酶。

（二）氯化铯纯化组织中的 RNA

纯化组织来源的 RNA 时要特别小心，因为某些组织（如胰腺、脾）中含有非常丰富的 RNA 酶（肝和肠的 RNA 酶水平相对低一些）。

① 迅速从动物体内取出组织冻存于液氮中。样品应切成重 2g 以下的组织块。组织中的 RNA 在体外非常不稳定，因此快速冻存组织非常重要。直接将组织放进胍盐溶液中等待碾碎会导致 RNA 降解。

② 约每 2g 组织加 20ml 组织胍盐溶液 [1L 溶液：在约 400ml DEPC 处理水中溶解 590.8g 异硫氰酸胍，加 25ml 2mol/L Tris-HCl（pH7.4）和 20ml 0.5mol/L Na_2EDTA（pH 8.0），搅拌过夜，定容至 950ml，过滤，最后加 50ml β-巯基乙醇，不含十二烷基肌氨酸钠]，立即用组织匀浆器完全碾碎组织。组织胍盐溶液中不含十二烷基肌氨酸钠很重要，否则会产生气泡。

③ 用角转头离心，10000r/min（12000×*g*），10min，12℃。

④ 回收上清液，加入 0.1 倍体积的 20% Sarkosyl，65℃加热 2min。

⑤ 每 1ml 样品液加 0.1g CsCl，待 CsCl 溶解后，将样品细心加在装有 9ml 5.7mol/L CsCl 的 SW-28 硅化并经高压消毒的同质异晶聚合离心管顶部，在 SW-28 转头中以 25000r/min(113000×*g*)，22℃离心过夜。

⑥ 按（一）方案中步骤⑥描述的方法小心除去上清液，倒置试管让残余液体流出。切下试管底部（含 RNA 沉淀）并放入 50ml 塑料试管中。

⑦ 加 3ml 组织重悬缓冲液（5mmol/L EDTA，0.5%十二烷基肌氨酸钠，5% β-巯基乙醇），4℃放置过夜，使 RNA 重悬。重悬 RNA 沉淀非常困难，有时样品需放置很长时间。重悬过程中高浓度的 β-巯基乙醇和十二烷基肌氨酸钠可防止 RNA 降解。

⑧ 先后用 25∶24∶1（体积比）酚-氯仿-异戊醇和 24∶1（体积比）氯仿-异戊醇抽提 2 次。

⑨ 加 1/10 体积 3mol/L 乙酸钠（pH5.2）和 3 倍体积无水乙醇沉淀 RNA，再用水重悬。按（一）方案步骤⑩进行定量。－70℃贮存。

（三）一步法从培养细胞或组织中分离 RNA

培养细胞或组织在含 4mol/L 异硫氰酸胍的变性液中进行匀浆，匀浆液依次与 2mol/L 乙酸钠（pH4.0）、酚-氯仿-异戊醇混合，混合液离心后的上清水相中含有总 RNA，而交界面和有机相中分别含蛋白质和 DNA。RNA 经异丙醇沉淀并重新溶解于变性液（含 4mol/L 异硫氰酸胍）中，异丙醇沉淀，75%乙醇洗涤。

以下步骤除特别指出外，均在室温下进行。

①a 处理组织时，每 100mg 组织加 1ml 变性液［4mol/L 异硫氰酸胍，25mmol/L 柠檬酸钠（pH7.0），0.1mol/L β-巯基乙醇，0.5%十二烷基肌氨酸钠。先配制贮存液：293ml 水、17.6ml 0.75mol/L 柠檬酸钠（pH7.0）和 26.4ml 10%十二烷基肌氨酸钠溶液，加入 250g 异硫氰酸胍，60～65℃搅拌溶解。贮存液室温下可存放 3 个月。使用时取 50ml 贮存液，加 0.35ml β-巯基乙醇即为工作液，这种变性工作液可在室温下存放 1 个月］，用玻璃特夫隆（Teflon）匀浆器撞击数次匀浆。

①b 处理培养细胞，对悬浮培养细胞离心后除去上清；对单层培养细胞，吸除培养液。每 10^7 个细胞加 1ml 变性液，用吸管上下抽吸裂解液 7～10 次。

注：不要用盐水洗涤细胞，单层培养细胞可以直接在培养皿或培养瓶中进行裂解。以下操作使用一次性灭菌、圆底、有盖聚丙烯试管，试管不需进行其他处理，但使用前应检测试管带有变性液和酚-氯仿时能否耐受 10000×*g* 离心。

② 将匀浆液转入 5ml 聚丙烯试管中，加 0.1ml 2mol/L 乙酸钠（在 40ml 水和 35ml 冰醋酸中加 16.42g 无水乙酸钠，冰醋酸调节 pH 到 4.0，最后用水定容至 100ml。最终溶液的钠离子浓度为 2mol/L)，上下颠倒充分混合。加 1ml 水饱和酚，充分混合，再加 0.2ml 49∶1（体积比）氯仿-异戊醇，混合充分后，0～4℃放置 15min。

注：混合时注意旋紧盖子。上面所给出的溶液体积均为每 1ml 变性液用量，如果开始所加的变性液容量不同，则随后的其他溶液容量也要相应变化。

③ 使用 SS-34 转头，4℃，9000r/min(10000×*g*) 离心 10min，将上层水相转入一新试管中。上层水相中含有 RNA，而 DNA 和蛋白质分别在界面和下层有机相中。水相体积约 1ml，与开始所用变性液体积相同。

④ 加 1ml（1 倍体积）100%异丙醇，10000×*g* 4℃离心 10min，弃上清。

注：从含高浓度糖原的组织（如肝脏）中分离 RNA 时，建议使用改良的一步法方案以减低糖原污染。

在异丙醇沉淀步骤后，加入 4mol/L LiCl，剧烈振荡将糖原从 RNA 沉淀中洗出，5000×g 离心 10min，沉淀不溶的 RNA，然后将 RNA 重新溶解于变性液中进行后面的操作。

⑤ 加 0.3ml 变性液溶解 RNA 沉淀，再转移至 1.5ml 微量离心管中。

⑥ 加 0.3ml（1 倍体积）100%异丙醇，置－20℃ 30min 沉淀 RNA；然后 4℃离心，10000×g 10min，弃上清。

⑦ 重悬 RNA 沉淀于 75%乙醇中，振荡，脱盐。置－20℃ 30min 沉淀 RNA，然后室温下温育 10～15min，以溶解残存的胍盐。

⑧ 10000×g 离心 5min，弃上清，真空干燥 RNA 沉淀 5～15min。

⑨ 依据以后 RNA 用途的不同，溶解 RNA 沉淀于 100～200μl DEPC 处理水或 DEPC 处理的 0.5% SDS 中。按（一）方案步骤⑩进行定量，－70℃贮存或在乙醇中－20℃贮存。

进行任何制备 RNA 的操作，都必须特别小心并确保所使用的溶液中无 RNA 酶。在加入胍盐溶液后除 TES（Tris 灭活 DEPC）外，所有接触 RNA 的溶液都必须用 DEPC 处理。

这里介绍的前两种方案都是根据梯度离心来完全分离 RNA 与 DNA 和蛋白质，其要点是使用硅化试管以及在吸取上清时仔细操作。RNA 产量较低的原因通常是离心后重悬 RNA 的时间不充分，RNA 沉淀溶解有一定困难，因此，应有足够的时间并充分振荡，使 RNA 完全溶解。

使用一步法分离 RNA 时应注意两点。首先，尽可能使用新鲜的组织，否则应立即将组织冻于液氮中并贮存于－70℃。后一种情况在处理时应先将组织在液氮中研成粉末，然后加入变性液，在融化前匀浆。第二，不要让最后形成的 RNA 沉淀完全干燥，这一点也非常重要，否则其可溶性会大大降价，这一点在所有的 RNA 分离方案中都同样重要。部分溶解的 RNA 其 A_{260}/A_{280} 比值为 1.6。通过在 55～60℃ 加热和间断振摇，或用移液器吸头抽吸 RNA 溶液可以增加 RNA 的溶解度。

使用（一）方案每 10^8 个细胞可得到约 200μg RNA，获得的 RNA 应完全无 DNA 污染。使用一步法分离得到的是所有各种片段的 RNA 分子，包括小分子 RNA（4S～5S），其产量根据分离的组织来源不同而有变化，每 100mg 肌肉组织约可获得 100～120μg。分离的 RNA A_{260}/A_{280} 比值应在 1.8 以上。

三、细菌 RNA 的制备

从细菌中分离 RNA 的过程包括裂解细胞和随后分离 RNA 并除去污染的 DNA 和蛋白质。下面介绍的方法裂解策略各有不同，如在革兰阴性菌是用蔗糖/去污剂或溶菌酶进行化学降解，而在革兰阳性菌则是用超声破碎胞壁。而后采用酶降解、有机溶剂抽提和乙醇或盐沉淀的方法将 RNA 与其他细胞成分分离，同时使用一些 RNA 酶活性抑制剂［如 DEPC、氧钒核苷（Vandy-ribonucleoside）复合物和金精三羧酸（aurin-tricarboxylic acid）］。如果需要极高质量的 RNA，下述（一）方案中采用的氯化铯梯度离心方法就可去除所有痕量污染 DNA。注意：所用的水和所有其他溶液均应用 DEPC 处理，以抑制 RNA 酶活性。

（一）从革兰阴性菌中分离高质量 RNA

本方案可以从大肠杆菌或蓝藻目细菌中获得引物扩增用的高质量 RNA，适于 Northern 印迹和 S1 核酸酶作图。首先用蔗糖/TritionX-100 溶液裂解细胞，而后用有机溶剂抽提和乙醇沉淀得到总核酸，最后经氯化铯梯度离心过程去除污染 DNA 和蛋白质，得到高质量 RNA。使用 Beckman TL-100 超速离心机（TLA-100.3 转头）可减少 RNA 沉淀时间，但许

多实验室没有配该转头，因此本方案也给出了使用 SW-41 转头离心的相应条件。此外，本方案还需要两种 RNA 酶抑制剂——氧钒核苷（VRC）复合物和金精三羧酸，不需要用溶菌酶和蛋白酶进行消化。

1. 提取核酸

① 培养 100ml 大肠杆菌或 500ml 蓝藻目细菌至对数生长期，加 1/20 体积终止缓冲液[200mmol/L Tris-HCl（pH 8.0），20mmol/L EDTA，20mmol/L NP-40，20mmol/L ATA]。终止其生长，将培养物置于冰上。可以直接把冰块加到培养液里进行降温。终止缓冲液中含有核酸酶抑制剂 ATA，此抑制剂可影响一些特定酶类，如果分离的 RNA 是为了进行引物延伸或 S1 核酸酶分析，则不应使用。

② 用 Beckman JA-14 转头，在 4℃以 6000r/min（5500×g）离心 5min 收集细胞。2ml STET 裂解液［8%蔗糖，5%（体积分数）Triton X-100，50mmol/L EDTA，50mmol/L Tris-HCl（pH7.0），用 DEPC 处理的贮备液制备，4℃贮存］重悬沉淀，加入 100μl 200mmol/L RVC，然后转至 15ml 聚丙试管中。

③ 加 1ml Tris 饱和酚，振摇 1min 后加 1ml 氯仿，再振摇 1min。用 JA-17 转头，4℃以 8500r/min（10000×g）离心 10min，沉淀用 2ml 10mmol/L RVC（ribonucleoside vanadyl complex，氧钒核糖核苷复合物）重悬。离心后，细胞碎片在界面处形成一厚层，收集上层水相时注意不要打动界面。

④ 加 1/10 体积 3mol/L 乙酸钠和 2 体积预冷无水乙醇沉淀核酸。用 JA-17 转头，在 4℃以 8500r/min（10000×g）离心 10min，沉淀用 2ml 10mmol/L RVC 重悬。加入冷乙醇后，核酸会立即形成可见沉淀物，不需要把试管置于低温下，核酸沉淀的形成提示细胞裂解完全。

⑤ 用酚-氯仿（1∶1，体积比）抽提 2 次，再按步骤④沉淀核酸。此步骤主要是降低蛋白质污染，沉淀后核酸体积应小于步骤④所见。

2. CsCl 梯度离心纯化 RNA

①a 使用 TLA-100.3 转头　在 13mm×51mm TLA-100.3 多聚碳酸盐（polycarbonate）管中加 0.75ml 5.7mol/L CsCl，同时用 2ml DEPC 处理水重悬核酸沉淀，加入 1g 固体 CsCl 并使其充分溶解，然后将此 2.25ml 溶液加到 CsCl 垫（5.7mol/L CsCl）上。

①b 使用 SW-41 转头　在 14mm×89mm ultraclear 离心管中加 3ml CsCl 垫，同时用 6ml DEPC 处理水重悬核酸沉淀，再加入 4.5g 固体 CsCl，用 DEPC 处理水补足体积到 9ml，然后将此溶液加到 CsCl 垫上。

不论使用何种转头，重要的是保证分离液体加到 CsCl 垫上时动作要轻柔。也可以使用其他吊桶式超速离心机，但样品体积、转速和离心时间都要进行相应调整。

②a 使用 TLA-100.3 转头　80000r/min（280000×g）离心 1h，20℃。

②b 使用 SW-41 转头　30000r/min（15000×g）离心 24h，20℃。

3. 回收 RNA

① 离心后立即用无菌巴斯德吸管从界面小心地吸除 DNA，接着再吸除上层 CsCl 垫，然后倒掉残余上清液，在 RNA 沉淀处作一标记，用薄纸擦净离心管管壁。注意离心后要立即开始下面操作，否则时间长后 RNA 会变松散，使从沉淀中除掉液体的操作出现困难。

② 用 0.36ml DEPC 处理水重悬 RNA 沉淀，使用 1000μl 移液器转至 1.5ml 微量离心管中。

③ 加 1/10 体积 3mol/L 乙酸钠和 2.5 体积预冷无水乙醇，—70℃沉淀 20min，4℃离心 5min 沉淀 RNA。

④ 加 1ml 预冷的 70%乙醇，4℃离心 5min。

⑤ 空气干燥 RNA，溶于 200μl DEPC 处理水中。测 A_{260} 和 A_{280} 定量，调整终浓度为 4μg/μl。—70℃长期贮存或以乙醇沉淀的形式—20℃贮存。

（二）革兰阳性菌 RNA 的分离

本方案用于从革兰阳性菌中分离 RNA，首先是超声破碎细胞壁，随后用去污剂裂解细胞膜以及蛋白酶消化降解细胞蛋白质，有机溶剂抽提和乙醇沉淀获得总核酸，酶去除 DNA，纯化 RNA。

① 4℃以 10000r/min（12000×*g*）离心 10min（Sorvall SS-34 转头），收集 10ml 细菌培养液。

② 重悬细胞于 0.5ml 裂解缓冲液［30mmol/L Tris-HCl（pH7.4），100mmol/LNaCl，5mmol/LEDTA，1%SDS，室温下贮存。使用前加蛋白酶 K 至浓度为 100μg/ml］中，转至微量离心管，在干冰上冰冻。

③ 融化细胞，超声破碎仪处理 3 次共 10s，功率设定约 30W。裂解充分后，细胞悬液应清亮，必要时可再次超声处理。注意避免裂解液起泡沫。

④ 37℃温育 60min，消化细菌蛋白质。

⑤ 加等体积酚-氯仿-异戊醇（25∶24∶1，体积比），室温下 10000×*g* 离心 5min，回收上层水相到干净微量离心管中。

⑥ 用等体积酚-氯-异戊醇（25∶24∶1，体积比）再抽提 1 次，然后再用等体积氯仿-异戊醇（24∶1，体积比）抽提。

⑦ 每 400μl 水相加 15μl 5mol/L NaCl，再在离心管中加满预冷无水乙醇，混合后冰浴 15～30min 或置于—20℃过夜。

⑧ 4℃ 10000×*g* 离心 15min 沉淀 RNA，用 500μl 预冷 70%乙醇洗涤沉淀物，然后在空气中晾干。

⑨ 用 95μl DNA 酶消化缓冲液［20mmol/L Tris-HCl（pH8.0），10mmol/L NaCl，1mmol/L 柠檬酸钠，1.5%SDS，室温下贮存］重悬沉淀，加 4μl 2.5mg/ml 无 RNA 酶的 DNA 酶Ⅰ，37℃温育 60min。

⑩ 用酚-氯仿-异戊醇（25∶24∶1，体积比）抽提，回收水相，在剩下的有机相中加 100μl TE 缓冲液，充分混合后室温 10000×*g* 离心 5min，回收并集中两次水相。

⑪ 再用氯仿-异戊醇（24∶1，体积比）抽提。

⑫ 在水相加 10μl 5mol/L NaCl，混合后再加 600μl 无水乙醇，—20℃沉淀过夜或在干冰-乙醇上沉淀 15min，在 4℃ 10000×*g* 离心 15～30min 收集 RNA 沉淀。

⑬ 用 500μl 预冷 70%乙醇洗涤沉淀，然后在空气中晾干，用 100μl DEPC 处理水重新溶解 RNA。取 10μl 稀释到 1ml，测 A_{260} 和 A_{280} 定量，—70℃贮存。

（三）革兰阴性菌 RNA 的快速分离

本方案步骤较简单，能很好地用于大肠杆菌或其他革兰阴性菌 RNA 的分离。先用溶菌酶破坏细胞壁，再用去污剂裂解余下的原生质体，裂解液中加有 RNA 酶抑制剂 DEPC。然后加入盐，共沉淀去污剂、蛋白质和染色体 DNA，并经离心除去，上清液中的主要成分为 RNA，经乙醇沉淀回收。用这种方法制备的 RNA 含有少量 DNA 和蛋白质，但是可用于大

部分 RNA 分析实验。

① 在 4℃以 10000r/min（12000×*g*）离心 10min，收集 10ml 革兰阴性菌。

② 重悬细菌于 10ml 原生质体缓冲液［15mmol/L Tris-HCl（pH 8.0），0.45mmol/L 蔗糖，8mmol/L EDTA，1.5% SDS，4℃贮存］中，加 80μl 50mg/ml 溶菌酶，冰浴 15min。溶菌酶裂解细胞壁，留下原生质体。

③ 在 4℃以 7000r/min（5900×*g*）离心 5min 收集原生质体。原生质体比较脆弱，因此使用较温和的离心过程。

④ 重悬于 0.5ml 革兰阴性菌裂解缓冲液［10mmol/L Tris-HCl（pH8.0），10mmol/L NaCl，1mmol/L 柠檬酸钠，1.5% SDS，室温下贮存］，加 15μl DEPC，转至微量离心管中。此时裂解液应变清亮和黏稠，避免过度振摇打断 DNA。

⑤ 37℃温育 5min。

⑥ 冰上冷却，加 250μl 饱和 NaCl（100ml DEPC 处理水中溶解 40g NaCl，搅拌直到溶液达饱和），上下颠倒混合，应形成大量沉淀物。沉淀中含 SDS、蛋白质和 DNA。

⑦ 冰浴 10min，室温或 4℃10000×*g* 离心 10min。

⑧ 转移上清液到两个干净微量离心管中，每管加 1ml 预冷无水乙醇，置于干冰上 30min 或－20℃过夜。

⑨ 4℃10000×*g* 离心 15min。

⑩ 用 500μl 70%冷乙醇洗涤沉淀然后在空气中晾干，用 100μl DEPC 处理水重新溶解 RNA。定量，－70℃贮存。

四、poly(A)$^+$RNA 的制备

本方案用于从带有大量 rRNA 和 tRNA 的总 RNA 中分离出含 poly(A) 的 RNA，首先对总 RNA 进行变性处理以暴露出 poly(A) 尾，然后带 poly(A) 的 RNA 与寡聚（dT）纤维素结合，其他 RNA 则被洗脱下来。层析柱子可以在去掉 poly(A)$^+$RNA 后重复使用。

1. 制备 oligo(dT) 层析柱子

① 用 10ml 5mol/L NaOH 洗涤硅化的柱子，并用水洗涤 1 次。可以用 2ml 的一次性柱子；或用硅化过的巴斯德吸管，头端用硅化过玻璃棉塞住。

② 在 1ml 0.1mol/L NaOH 溶液中加入 0.5ml oligo(dT) 纤维素干粉，将混合后的黏稠物加到柱子中，然后用约 10ml 水过柱洗涤。

③ 用 10～20ml 加样缓冲液［0.15mol/L LiCl，10mmol/L Tris-HCl（pH7.5），1mmol/L EDTA，0.1%SDS］平衡柱子，最后流出物的 pH 应为 7.5 左右。

2. 分离 poly(A) RNA

① 取约 2mg 总 RNA，70℃水浴 10min，加入 10mol/L LiCi 贮存液至终浓度为 0.5mol/L。

注：加热 RNA 是为了打断可能形成的二级结构。对所选择分离的 RNA 而言，采用的柱床体积不能过大，否则最终得到的 poly(A)$^+$RNA 稀释度较高，而使样品的处理和沉淀效率低下。因此选择分离的 poly(A)$^+$RNA 在 500μg 以下时应采用更小体积的层析柱，并相应减少各步骤中溶液的用量。通常，1ml oligo(dT) 纤维素足以分离 5～10mg 总 RNA。

② 将 RNA 溶液装柱，用 1ml 加样缓冲液洗涤层析柱，从加样开始收集所有的流出液。

③ 将收集到的流出液重复过柱 2 遍。最初的 RNA 共过柱 3 遍，以保证所有的 poly

(A)$^+$RNA 均与 oligo(dT) 结合。

④ 用 2ml 洗涤缓冲液 [0.15mol/L LiCl，10mmol/L Tris-HCl (pH7.5)，1mmol/L EDTA，0.1% SDS] 洗脱 2 遍。

⑤ 用 2ml 2mol/L EDTA-0.1% SDS 洗脱 RNA，流出液用新离心管回收。

⑥ 用加样缓冲液重新平衡 oligo(dT) 柱子，然后将 RNA 流出液按步骤①～⑤重复 1 次。上述重复操作步骤可以进一步纯化带 poly(A) 尾的 mRNA，如果对 mRNA 的要求不是特别高时，此步骤可以省略。

⑦ 用 3mol/L 乙酸钠贮存液调整 RNA 流出物的乙酸钠浓度至 0.3mol/L，加 2.5 倍体积的无水乙醇，然后转至两个硅化过的 SW-55 离心管中。

⑧ −20℃沉淀过夜或放置在干冰-乙醇上沉淀 30min，用 Beckmam SW-55 转头在 4℃以 50000r/min 离心 30min，收集沉淀物。

⑨ 倒出乙醇，沉淀置于空气中干燥，然后用 150μl 无 RNase 的 TE 缓冲液重悬沉淀。

绝大多数 mRNA 都带有 poly(A) 尾，而 tRNA 和 rRNA 则没有，因此经 poly(A) 选择后可以富集 mRNA。用此方法进行 cDNA 文库的构建已证实其有效，而在分析低丰度 mRNA 的结构时，此方法也非常有效。在 RNA 的总制备物中去除 rRNA 和 tRNA 后可以得到的 mRNA 量很高，采用后者进行 S1 核酸酶分析就可以检测到低水平的信息。

在进行 poly(A) 选择操作时必须使用无 RNA 酶的溶液，要求比其他有关 RNA 的操作还要严格，这是由于多数情况下 mRNA 的 5′端是必需的，如果在 5′端和 3′端之间发生断裂，则 mRNA 将会与其 poly(A) 尾分离。另外一个需要注意的问题是柱子的体积必须与所选择的 RNA 量相配合，柱子的体积大小由洗出的 poly(A)$^+$RNA 的量决定，如果柱子体积过大，那么洗出的 poly(A)$^+$RNA 会被稀释，RNA 越稀释，沉淀越困难。此外，由于稀释的 RNA 溶液可非特异地黏附到柱子和沉淀时使用的离心管管壁上，会造成较大量丢失。oligo(dT) 纤维素通常由制造商提供，容量和结合能力很大，当从小量 RNA 中进行 poly (A)$^+$RNA 分离时，小体积柱子就已经足够，而洗脱液体积也应相应减少。RNA 中约 1% 为 poly(A)$^+$RNA，在琼脂糖凝胶电泳上成为一条纤细的带，大小约 20kb（如密度很高，也可能为 5～10kb 大小），同时不应看到 rRNA 带。

第四节　核酸的定量测定

核酸的定量方法有多种，较为常用的主要是紫外分光光度法和电泳比较法，另外也可采用化学检测的方法，如定磷法、二苯胺法和地衣酚法等。

一、紫外分光光度法测定 DNA 和 RNA 的含量

本法要求核酸样品是纯净的（即无显著的蛋白质、酚、琼脂糖或其他核酸、核苷酸等污染）。用紫外分光光度计测定 260nm 和 280nm 两个波长处的光吸收，260nm 和 280nm 两处读数的比值（A_{260}/A_{280}）可反映核酸的纯度。DNA 和 RNA 纯品的 A_{260}/A_{280} 值分别为 1.8 和 2.0，如果样品中有蛋白质或酚的污染，则 A_{260}/A_{280} 将明显低于此值，此时就无法对样

品中的核酸进行精确定量，可将样品纯化后再作定量测定。

组成核酸分子的碱基均具有一定的吸收紫外线特性，最大吸收值在250～270nm之间。例如腺嘌呤的最大紫外线吸收值在260.5nm，胞嘧啶在267nm，鸟嘌呤在276nm，胸腺嘧啶在264.5nm，尿嘧啶在259nm。这些碱基与戊糖、磷酸形成核苷酸后，其最大吸收峰基本不会改变，核酸的最大吸收波长是260nm，吸收低谷在230nm，这个物理特性为测定核酸溶液浓度提供了基础。

在260nm波长紫外光下，*A*值为1时的吸光度相当于双链DNA浓度为50μg/ml，相当于单链DNA或RNA为40μg/ml，相当于单链寡核苷酸为20μg/ml，可以由此计算核酸样品的浓度。

二、电泳比较法（溴化乙锭荧光法）

核酸样品可通过琼脂糖凝胶电泳分离后，用荧光染料染色的方法确定样品的含量。这种方法既适用于DNA的检测，也同样适合RNA样品的检测。两者的检测方法一致，但由于RNA样品通常不稳定，容易降解，因此该方法更多用于DNA的电泳。对于RNA电泳，应当采用专用电泳槽，并需要进行较为严格的抑制RNA酶（RNase）处理，如用DEPC水处理及添加RNA酶抑制剂等，以确保RNA样品在电泳过程中不会被降解。

通常将质粒DNA或微量DNA（<250ng/ml）样品与已知浓度的DNA标准同时电泳，根据样品DNA和标准DNA结合荧光染料（如溴化乙锭）的多少，在紫外线照射下观察比较样品与标准的荧光强度，或用凝胶分析软件分析，确定样品的浓度。

如果DNA样品的纯度不高，其中的杂质可能参与荧光发射，也可能猝灭荧光。为解决这个问题，可将待测DNA样品与一系列已知不同浓度的DNA标准分别取2μl，加入0.4μl电泳上样缓冲液（含溴酚蓝），混匀后加样于含0.5μg/ml溴化乙锭的琼脂糖微型电泳凝胶的不同加样孔内，同时进行电泳，使DNA移入琼脂糖胶内，一般在溴酚蓝迁移2cm左右可结束电泳。将凝胶浸入含0.01mol/L $MgCl_2$的电泳缓冲液中5min，进行背景脱色，然后即可在短波紫外线（254nm）下进行观察或拍照，比较样品DNA与DNA标准的荧光强度，即可大致判断出待测样品中DNA浓度。

溴化乙锭荧光法灵敏、快速，但它的荧光强度决定于溴化乙锭嵌入碱基中数量的多少，这又与DNA的超螺旋程度密切相关，样品与标准往往很难完全一致，所以此法只能用于估算，不能进行精确定量。应当注意：溴化乙锭是强致癌物，注意防护和控制污染。近年来许多公司开发了溴化乙锭的替代产品，可作为DNA电泳凝胶染色的新型荧光染料，染色效果良好。

三、定磷法测定核酸

核酸中含有磷酸，其比例相对较为恒定，一般DNA含磷量为9.9%，RNA含磷量约为9.5%。因此可以根据样品中磷的含量计算核酸（DNA或RNA）的含量。测定时，样品用强酸（5mol/L硫酸或60%过氯酸）消化，使核酸中含的磷转变为无机磷，后者在酸性条件下与钼酸铵反应生成磷钼酸。在还原剂的作用下，磷钼酸生成蓝色的化合物——钼蓝。

钼蓝的最大吸收峰在650～660nm波长处，可通过测定650nm处的吸光度值，确定磷的含量，进一步计算出核酸的含量。操作要点如下。

1. 样品的消化

取 1ml 1%的核酸溶液，加入 2～3ml 5mol/L 硫酸于消化管中，调节温度在 140～160℃，消化 2～4h，待溶液呈黄色后取出，稍冷后加入 1～2 滴 30%的过氧化氢，继续消化至溶液透明为止。取出冷却后加 1ml 蒸馏水，在沸水浴中加热 10min，以分解消化过程中出现的焦磷酸。然后移入 100ml 容量瓶中定容至刻度。

2. 磷的测定

取上述定容的样品液 1ml，加蒸馏水 2ml 和定磷试剂（3mol/L 硫酸：水：2.5%钼酸铵：10%维生素 C=1：2：1：1，当天混合使用）3ml，混匀后放置 45℃水浴中反应 25min。冷却至室温后，于 650nm 波长处测定吸光度。减去空白值就可以从标准曲线中查出无机磷的含量。

标准曲线的制作：以不同含量的无机磷溶液代替样品液，按上述方法分别测定 650nm 处吸光度值，然后以标准磷含量（μg）为横坐标，吸光度为纵坐标，绘制标准曲线。

最后参考下列公式计算核酸量，计算公式为：

$$\text{DNA 量}=(\text{总磷量}-\text{无机磷量})\times 10.9$$

$$\text{RNA 量}=(\text{总磷量}-\text{无机磷量})\times 10.5$$

四、二苯胺法测定 DNA 含量

DNA 中的 2-脱氧核糖在酸性条件下与二苯胺试剂一起加热产生蓝色反应，在 595nm 处有最大吸收峰。在 40～400μg 范围内，A_{595} 与 DNA 的浓度成正比。操作要点如下。

1. 配制二苯胺试剂

取 1.5g 二苯胺，溶于 100ml 冰醋酸中，添加 1.5ml 硫酸，暗处保存备用。使用前加入 1ml 1.6%乙醛溶液（少量乙醛可提高反应灵敏度）。

2. 制作标准曲线

取一定量的标准 DNA，用 0.01mol/L 的氢氧化钠溶液配置成 200μg/ml 的 DNA 标准液。分别吸取 0、0.4ml、0.8ml、1.2ml、1.6ml、2.0ml 的 DNA 标准液，用 0.01mol/L NaOH 溶液补充至 2ml。各加入 3ml 二苯胺试剂，混匀，于 60℃水浴中保温 1h。冷却后测定 A_{595}。以 DNA 浓度为横坐标，吸光度为纵坐标，绘制标准曲线。

3. 样品中 DNA 含量的测定

准确称取一定量的样品，用 0.01mol/L NaOH 溶液配成一定浓度，取 2ml 样品液，加入 3ml 二苯胺试剂，按上述方法测定 A_{595}。然后从标准曲线查出样品中的 DNA 含量。计算样品中 DNA 的含量。

五、地衣酚法测定 RNA 含量

核糖核酸在浓盐酸和三氯化铁存在下，与地衣酚（3,5-二羟甲苯）反应，生成绿色物质，其最大吸光度在 670nm 波长处。在一定浓度范围内，吸光度与 RNA 浓度成正比。此方法测定时存在 DNA 干扰，评价结果时要考虑。操作要点如下。

1. 配制地衣酚试剂

0.1g $FeCl_3 \cdot 6H_2O$ 溶于 100ml 浓盐酸中，使用前加入 0.1g 地衣酚。

2. 制作标准曲线

称取标准 RNA，用蒸馏水配成 100μg/ml 的 RNA 标准液。分别吸取 0、0.5ml、

1.0ml、1.5ml、2.0ml、2.5ml 的 RNA 标准液，用蒸馏水补充至 2.5ml，分别加入 2.5ml 地衣酚试剂。混匀后，于沸水浴中加热 30min，冷却后测定 A_{670}。以 RNA 浓度为横坐标，A_{670} 为纵坐标，绘制标准曲线。

3. 样品的测定

取 2.5ml 样品液，加入 2.5ml 地衣酚试剂，测定 A_{670}，从标准曲线中查出样品液中 RNA 的含量，再算出样品中 RNA 的含量。

（尹燕华　屈伸）

第三章　PCR 技术

PCR（polymerase chain reaction）技术即聚合酶链反应，又称 DNA 体外扩增技术，是 1985 年由美国 PE-Cetus 公司的人类遗传研究室 Kary Mullis 等人发明的一种具有划时代意义的分子生物学技术。此技术可以在生物体外，在几个小时的时间内将极微量的目的基因成百万倍地扩增。这一技术操作简单，容易掌握，结果也相对可靠，为基因的分析与研究提供了一种强有力的手段。它使一度非常稀有的实验所需遗传物质变得丰富，遗传物质不但总量变多了，而且不再受限于活的生物体，因此极大地提高了基因操作的效率及基因操作的灵活性。PCR 技术是方法学上的一次革命，以其显著的三大特点——特异性、高效率和忠实性，对生命科学研究领域产生了巨大的影响。PCR 被誉为无细胞分子克隆，由于 Mullis 的这一杰出贡献，这位年轻的科学家获得了 1993 年度诺贝尔化学奖。

第一节　PCR 技术基础

一、基本原理

PCR 是一种选择性体外扩增 DNA 的方法。其原理类似于细胞内发生的 DNA 复制，在试管中给 DNA 的体外合成提供一种合适条件——模板、dNTP、寡核苷酸引物、DNA 聚合酶、合适的缓冲液系统，以及 DNA 变性、复性及延伸的温度与时间等，就能够特异地扩增任何目的基因或 DNA 片段。

（一）反应过程

PCR 全过程包括三个基本步骤，即双链 DNA 模板加热变性成单链（变性）；在低温下引物与单链 DNA 互补配对（退火）；在适宜温度下 *Taq* DNA 聚合酶以单链 DNA 为模板，利用 4 种脱氧核苷三磷酸（dNTP）催化引物引导的 DNA 合成（延伸）。这三个基本步骤构成的循环重复进行，可以使特异性 DNA 扩增达到数百万倍（$\geqslant 2\times 10^6$）。该过程见图 3-1 所示。

（1）变性　加热至 90～96℃时，模板 DNA 双螺旋的氢键断裂，双链解链，形成单链 DNA。

（2）退火　当温度突然降低至 25～65℃，引物与互补的单链 DNA 模板在局部形成杂交链。

（3）延伸　70～74℃下，在 Mg^{2+} 存在的条件下，4 种脱氧核糖核苷三磷酸底物在 *Taq* DNA 聚合酶的作用下，引物沿 $5'\rightarrow 3'$ 方向延伸，按碱基互补原则合成与模板 DNA 互补的 DNA 新链。

变性、退火、延伸三个步骤称为 PCR 的一轮循环。每循环一次，目的 DNA 的拷贝数加倍，经过 n 次的循环后，PCR 扩增倍数为 $(1+X)^n$，X 为扩增效率，平均为 75%。循环

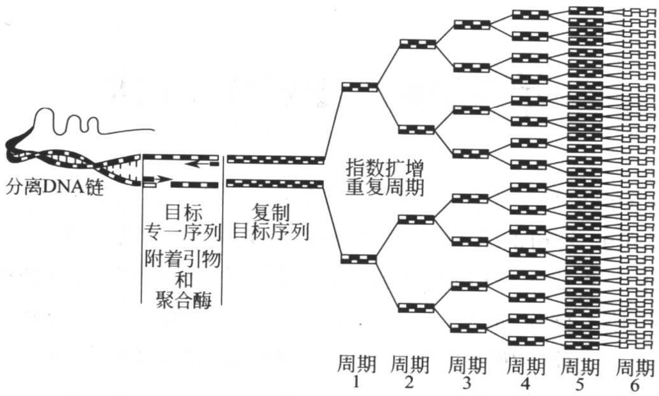


图 3-1　聚合酶链反应

次数一般为 25～30 次。如果一次循环需 2～3min，1～3h 就能将目的基因放大几百万倍。扩增反应可在 DNA 热循环仪中完成。

（二）产物类型

仔细分析 PCR 反应过程可以发现 PCR 反应产物并不完全一样，其中至少可以分为两类产物，即所谓“长片段产物”和“短片段产物”。它们是由引物所结合的模板不同造成的。长片段产物带有引物结合部位以外的序列，而短片段产物是严格限定在引物 5′端之间的序列，如图 3-2 所示。以一个双链 DNA 分子的原始模板为例，其变性、退火后，引物可结合到模板中与引物互补的序列，经延伸反应后得到双链产物。由于延伸反应总是从 5′→3′端，因此，延伸反应的终止点取决于模板的长度。在第 1 至第 2 个 PCR 循环中均是长片段产物，然而从第 3 个循环开始，产物中出现了限定在 2 个引物之间的短片段模板。显然，短片段是指数倍数扩增，而长片段是以算术级数增加，几乎可以忽略不计。

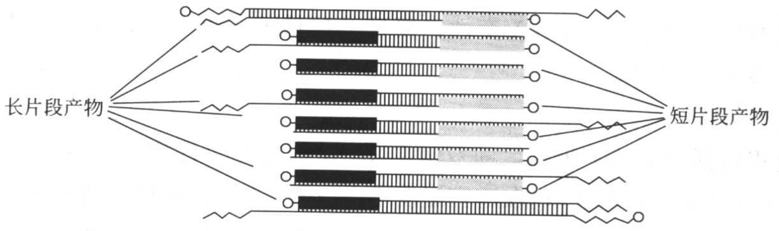


图 3-2　PCR 中的“长片段产物”和“短片段产物”
实体和阴影线段分别为两端的引物

（三）平台效应

PCR 扩增能力惊人，理论上 PCR 合成产物（短产物片段）的数量经过每轮循环将增加一倍，PCR 反应 30 轮循环后，PCR 扩增量应达到 2^{30} 个拷贝，约 10^9 个拷贝，但由于 DNA 聚合酶的质量、待扩增片段的序列及反应系统的条件等多种因素的影响，实际扩增效率比预期的要低，应按 $N_f=N_0(1+Y)^n$ 的指数方式递增（其中 Y 代表扩增反应效率，最大值为 1），一般可达 10^6～10^7 个拷贝。随着 PCR 循环次数的增加，合成产物达到 0.3～1pmol/L 水平，由于产物的堆积，dNTP、引物及 *Taq* DNA 聚合酶的消耗等原因，使产物原以指数增加的速率逐渐变成平坦曲线，即平台效应，如图 3-3 所示。这时扩增 DNA 片段的增加减

慢，进入相对稳定状态，再增加 PCR 的循环次数也不能增加目的 DNA 片段。到达平台期所需 PCR 循环次数取决于样品中模板的拷贝数、PCR 扩增效率、*Taq* DNA 聚合酶的活性及非特异性产物的竞争等因素。欲在达到“平台期”以前增加目的 DNA 片段的积累，必须尽量避免或减少非特异产物的产生。

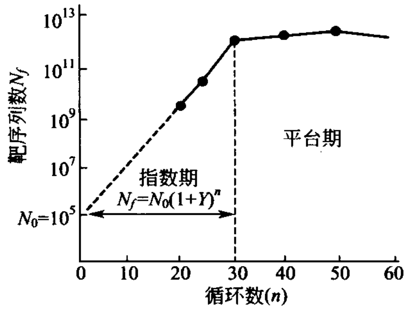


图 3-3　PCR 产物积累曲线

平台效应在 PCR 反应中是不可避免的，但一般在平台效应出现之前，合成的目的基因的数量已可满足实验的需要。例如，PCR 标准反应中加入 1μg 人类基因组 DNA 模板（一般从 1ml 血样中可以获得 50μg DNA)。它应含有各种单拷贝序列分子 3×10^5 个，相当于 0.1pg（或 10^{-13}g）的 300 个核苷酸长的靶 DNA 序列。经过 30 次循环扩增之后，便可产生出约 1μg 的 DNA 产物片段，足以满足任何常用的分子生物学研究，包括直接的 DNA 序列测定的需要。有报道，即便反应混合物中只含有一个拷贝的靶 DNA 分子，亦能被有效地扩增。正因为 PCR 技术具有如此高的扩增敏感性，所以任何 DNA 样品或实验试剂均应避免污染。同时，也意味着这一技术在痕量 DNA 的样品检测鉴定领域（如法医学中鉴定一根头发、一丝血迹等样品）具有重要的应用价值。

二、PCR 技术的特点

1. 特异性强

PCR 反应的特异性决定因素为：①引物与模板 DNA 特异正确的结合；②碱基配对原则；③*Taq* DNA 聚合酶合成反应的忠实性；④靶基因的特异性与保守性。其中引物与模板的正确结合是关键，它取决于所设计引物的特异性及退火温度。在引物确定的条件下，PCR 退火温度越高，扩增的特异性越好。*Taq* DNA 聚合酶的耐高温性质使反应中引物能在较高的温度下与模板退火，从而大大增加 PCR 反应的特异性。

2. 灵敏度高

从 PCR 的原理可知，PCR 产物的生成是以指数方式增加的，即使按 75％的扩增效率计算，单拷贝基因经 25 次循环后，其基因拷贝数也在一百万倍以上，即可将极微量（pg 级）DNA，扩增到紫外光下可见的水平（μg 级）。

3. 简便快速

现已有多种类型的 PCR 自动扩增仪，只需把反应体系按一定比例混合，置于仪器上，反应便会按所输入的程序进行，整个 PCR 反应在数小时内就可完成。扩增产物的检测也比较简单：可用电泳分析，不用同位素，无放射性污染，易推广。

4. 对标本的纯度要求低

不需要分离病毒或细菌及培养细胞，DNA 粗制品及总 RNA 均可作为扩增模板。可直接用各种生物标本（如血液、体腔液、毛发、细胞、活组织等）粗制的 DNA 扩增检测。

三、标准 PCR 反应

1. 反应体系

标准 PCR 反应体积为 50～100μl，其中含有：1×PCR 反应缓冲液［50mmol/L KCl,

10mmol/L Tris-HCl（pH8.3）、1.5mmol/L $MgCl_2$，100μg/ml 明胶]、4 种 dNTP（dATP、dCTP、dGTP、dTTP）各 200μmol/L、2 种引物各 0.2～1μmol/L、DNA 模板 0.1μg（需根据具体情况加以调整，一般需 10^2～10^5 拷贝的 DNA）、*Taq* DNA 聚合酶 2 个单位、ddH_2O 补至终体积（50～100μl）。

现在各仪器公司生产的 PCR 仪可选用 0.2ml 或 0.5ml 的 PCR 管，前者更为常用，PCR 反应体系最小可达 5μl，各反应组分可按相应比例缩小。

2. 反应步骤

在 0.5ml 无菌 Eppendorf 管中，依次加入下列试剂：PCR 反应缓冲液、4 种 dNTP、引物、DNA 模板，混匀，95℃加热 10min，以除去 DNA 样品中蛋白酶、氯仿等对 *Taq* DNA 聚合酶的影响。然后，每管加入 2U *Taq* DNA 聚合酶，混匀，离心 30s，加 50μl 石蜡油封盖反应体系防止反应液的挥发，置离心管于 PCR 仪中开始 PCR 循环。

新型 PCR 仪设计了热盖，使反应中整个管温度均匀，盖与管底温度一致，避免了加热过程中液体在管盖的冷凝。

3. PCR 循环参数

依据设计的循环参数进行扩增：

变性	90～94℃	30～50s
退火	25～65℃	60～90s
延伸	70～74℃	60～120s

反复循环 25～35 次。末次循环后，延伸温度再延时 5min。

4. 结果检测

取 PCR 扩增产物 5～10μl，加 1～2μl 上样缓冲液，于适当浓度的琼脂糖凝胶中进行电泳分析。为确定扩增产物是否符合设计要求，需选择合适分子量 DNA 标准平行电泳。

四、影响因素

特异性、高效性和忠实性是检测 PCR 扩增效率的三个指标。高度特异的 PCR 反应只产生一个扩增产物；而扩增反应越有效，经过相对少的循环会产生更多的产物。为了得到理想的 PCR 扩增效果，必须注意以下几个方面。

（一）模板

PCR 对模板的要求不高，单、双链 DNA 均可，也可以是 RNA，后者的扩增需首先逆转录成 cDNA 后才能进行正常 PCR 循环。为保证结果的准确性和特异性，PCR 反应中的模板加入量一般为 10^2～10^5 个拷贝靶序列，即一般宜用 ng 级的克隆 DNA、μg 级的染色体 DNA 或 10^4 个拷贝的待扩增片段来做起始材料；人基因组 DNA 1μg 相当于 3×10^5 个单拷贝靶分子，大肠杆菌 DNA 1ng 相当于 3×10^5 个单拷贝靶分子。因此扩增不同拷贝数的靶序列时，加入的含靶序列的 DNA 量也不同。另有资料认为，用小分子量和线性模板 DNA 扩增效果较好。因此当使用极高分子量的 DNA（如基因组 DNA）时，可以使用低频率切点的限制酶（如 *Not* Ⅰ或 *Sat* Ⅰ）先行消化，再作扩增效果好。闭合环状质粒作 PCR 模板时最好先线性化处理，以提高扩增效率。

作为 PCR 模板的核酸标本来源广泛，用于提取核酸的材料种类可能差异很大，如 1000 多万年前的木兰科叶子化石、已灭绝动物皮中发现的肌肉、单根人类毛发和石蜡包埋的活组织标本等；也可以从培养的细胞和微生物中提取，从临床标本（血、尿、体腔积液等）及其

他标本（血斑、毛发、精斑等）中直接提取。虽然大多数 PCR 对模板的要求都不高，纯度要求也不严，用量也低，待扩增核酸仍需要部分纯化以除去核酸标本中的蛋白酶、核酸酶、*Taq* DNA 聚合酶抑制物以及能结合 DNA 的蛋白质。纯化核酸的目的主要在于：①除去杂质，特别是除去干扰 *Taq* 酶活性的物质；②使待扩增的靶序列 DNA 暴露和浓缩，从而保证有足量的 DNA 模板启动 PCR 反应；③有利于评价扩增体系的灵敏度并根据产物对靶 DNA 进行定量分析。

提取 DNA 的基本过程是利用蛋白酶 K 及 SDS，在 EDTA 存在的条件下，裂解细胞，消化蛋白质，使核蛋白解聚及胞内 DNA 酶失活；然后用酚、氯仿多次提取去除蛋白质。在 DNA 中若混有少量 RNA，可用 RNA 酶去除。最后用乙醇沉淀得到纯 DNA。在提取中应尽量保持 DNA 的完整性和纯度，防止 DNA 降解。被扩增的 DNA 特定序列不需要事先从样品 DNA 中分离，因为 PCR 产物的序列，也即反应的特异性是由寡核苷酸引物所决定的。

单、双链 DNA 和 RNA 都可作为 PCR 的模板，如果起始模板为 RNA，需先通过逆转录得到第一条 cDNA 链后才能进行 PCR 扩增。

（二）引物

引物是待扩增核酸片段两端的已知序列，它决定了 PCR 扩增产物的大小。引物的选择是整个 PCR 扩增反应成功的关键因素，引物的优劣直接关系到 PCR 的特异性及成功与否。理想的引物应该有效地与靶序列杂交，而与出现在模板中的其他相关序列的杂交是可以忽略的。

1. 引物合成的质量

PCR 所用引物质量较高，且需纯化。合成的引物必须经聚丙烯酰胺凝胶电泳或反向高效液相层析（HPLC）纯化，以减少发生非特异扩增和信号强度的降低，冻干引物可以在常温下运输。冻干引物于－20℃可以保存 12～24 个月，甚至更长，液体状态于－20℃可以保存 6 个月或更长。

2. 引物设计的一般原则

引物设计的目的是要达到两个目标，即理想的扩增特异性和扩增效率。正确的引物设计，可以遵循如下的一些主要原则。

（1）引物长度　特异性一般通过引物长度和退火温度控制。寡核苷酸引物长度一般为 15～30 碱基。这些寡核苷酸引物适于不含序列变异的、靶位确定的标准 PCR。引物长度增加，能提高特异性，但是在退火时被引发的模板会减少，而降低反应效率。

实际应用中最适延伸温度不要超过 *Taq* DNA 聚合酶的最适温度（74℃），这样才能保证产物的特异性。

（2）引物的 3′末端和 5′末端核苷酸　引物 3′末端是 PCR 延伸的起始端，不能进行任何修饰，也不能有形成二级结构的可能，一般 3′端也不能发生错配。

引物 5′端仅起限定 PCR 产物长度的作用，它对扩增特异性影响不大，可以被修饰而不影响扩增的特异性。

（3）GC 的合理含量　一对引物的 GC 含量和 T_m 值应该协调，GC 含量一般为 40%～60%，如含有 50% GC 的 20 个碱基长度的寡核苷酸链的 T_m 值大概在 56～62℃范围内，这样可为退火提供足够高的温度。根据公式 $T_m=4(G+C)+2(A+T)$，可估计引物 T_m 值。

（4）避免引物自身和引物之间的互补　引物自身和引物之间均应避免互补序列或互补

性，尤其应避免3′端的互补重叠。一般来讲，引物自身连续互补碱基不能超过3个，一对引物间也不宜多于4个连续碱基的互补性。

(5) 碱基的随机分布　选择缺乏连续单一核苷酸的区域作为引物序列，这样可以减少引物-引物同源互补的机会。同时，也要注意引物对在长度和碱基组成上的平衡，使两个引物的 T_m 值相差在2～3℃范围内。

(6) 避开产物的二级结构区　采用计算机软件可以预测估计 mRNA 的稳定二级结构，有助于选择模板。

3. 使用引物设计软件

现在有商品化的或免费的计算机软件程序，可用来在确定的区域中选择引物序列。在对所研究基因组全部序列并不完全知道的情况下，这些程序既方便又实用，但是即使是最好的程序也并非完美无缺。

使用计算机软件时，选择参数设置越广泛，计算机需要考虑的情况越多，寻找引物所需的时间会越长。因此，在实验设计上用来寻找引物的参数应尽可能狭窄和独特，这样可以得到快速和高质量设计的引物。

4. 引物浓度

一般 PCR 反应体系中引物的终浓度为0.2～1μmol/L，在此范围内 PCR 产物量基本相同。引物浓度低于0.2μmol/L 时，PCR 产物量降低，引物浓度太高又会促进引物的错误引导，导致非特异扩增，还会增加引物二聚体的形成。非特异产物和引物二聚体又可作为 PCR 反应的底物，与靶序列竞争 DNA 聚合酶和 dNTP 底物，从而使靶序列的扩增量降低。可采用系列稀释法确定引物的最低工作浓度。经 DNA 合成仪合成的引物，一般须经 PAGE 电泳纯化，紫外分光光度计定量。根据定量结果计算配制 50μmol/L 引物贮存液的简便公式如下：

$$V=\frac{20\times W}{M\times 10^{-3}}$$

式中　V——表示配制 50μmol/L 引物浓度时所加的体积（μl）；

W——合成引物的实际质量（μg，按 $OD_{260}\times 30$ 计算）；

M——引物的分子量（按引物的碱基数×320计算）。

按该公式计算的体积溶解引物后，溶液中引物的浓度即为 50μmol/L。使用时若反应体系中引物的浓度为 1μmol/L 时，则 50μl 总反应体系中只需加 1μl 贮存液即可。

（三）循环参数

1. 变性的时间和温度

变性步骤必须使靶基因模板完全变性，使 DNA 双螺旋的氢键断裂，双链解离，形成单链 DNA，扩增才会成功。不同的 DNA 解链温度不同，这是由于 DNA 中 GC 与 AT 的比例不同引起的。GC 之间有3个氢键连接，AT 之间有2个氢键连接，所以 GC 比例大，则解链温度高。一般地讲，GC 含量每增加1%，解链温度增加0.4℃。哺乳动物基因组中 DNA 含40% GC 时，$T_m\approx 87$℃，含60% GC 时，$T_m\approx 95$℃。典型的变性条件是95℃ 30s 或97℃ 15s，更高的温度解链会更彻底，尤其是对富含 GC 的靶基因，但可能使 *Taq* DNA 聚合酶很快失去活性。如果变性温度低于90℃，则 DNA 变性不完全。变性温度在90～95℃时，既能保证使 DNA 双链模板变性，又能保持 *Taq* DNA 聚合酶活力。因此 PCR 变性温度在90～95℃之间选择，时间为30～60s。

2. 退火的时间和温度

退火温度决定着PCR的特异性。引物复性所需的温度与时间取决于引物的碱基组成、长度。引物长度为15～25bp时，其退火温度一般应低于引物 T_m 值5℃左右。引物的 T_m 值可根据公式 $T_m=4(G+C)+2(A+T)$ 进行计算，退火温度的范围一般在25～65℃之间，增加退火温度可减少引物与模板的之间的非特异性结合，提高PCR反应的特异性；降低退火温度则可增加PCR反应的敏感性。在典型的引物浓度（如0.2μmol/L）时，退火时间一般为40～90s，时间过短会导致引物与模板DNA互补配对失败。

3. 延伸的时间和温度

延伸时间的长短取决于待扩增的模板序列的长度和浓度，以及延伸温度高低。引物延伸是在70～74℃下进行，在72℃时合成核苷酸的速率约为35～100个核苷酸/s，延伸时间根据扩增DNA的长度而确定，一般延伸30～90s。

4. 循环次数的确定

以既达到扩增效果又尽量减少非特异性产物的扩增为原则，PCR的循环次数一般在25～35次，循环次数过多将增加非特异产物，循环次数太少，则产率偏低。因此在得到足够产物的条件下应尽量减少循环次数。

（四）*Taq* DNA聚合酶及其浓度

PCR反应体系中 *Taq* DNA聚合酶用量为1～2.5U/100μl。然而，由于模板、引物以及其他条件的差别，聚合酶的使用量也有差异，可根据电泳的结果决定酶的最佳用量。若酶的浓度偏高，非特异性的产物将会增加；若酶的浓度偏低，则合成产物的量减少。

由于生产厂使用的配方、制造条件以及活性定义不同，不同厂商供应的 *Taq* 酶的性能有所不同。*Taq* 酶应具有较好的热稳定性，Cetus公司分析结果为92.5℃、95℃、97.5℃下 *Taq* 酶的半衰期分别为130min、40min和5～6min。PCR变性温度通常设为94℃，而变性时间通常设为30～60s，30个循环累计热变性时间约15～30min，此时该酶仍可保持相当高的活性，完全可保证PCR反应的需要。

Taq DNA聚合酶催化DNA合成的温度以70～80℃为宜。在75～80℃条件下，该酶的 K_{cat} 值约为150个核苷酸/(秒・酶分子)；70℃时为60个核苷酸/(秒・酶分子)；55℃为24个核苷酸/(秒・酶分子)；而高于90℃时，DNA合成几乎不能进行。

需要特别注意的是PCR扩增过程中存在有dNTP错误掺入的可能性，*Tag* 酶错配的概率约为2/10000。而 *Taq* 酶没有 $3'\rightarrow5'$ 外切核酸酶活性，所以如果发生dNTP的错误掺入时，这种酶没有校正能力，任何错误都将保留在最后的结果中，这在利用PCR产物进行DNA序列分析时需要引起高度的注意。

（五）Mg^{2+} 浓度

Taq DNA聚合酶是 Mg^{2+} 依赖性酶，对 Mg^{2+} 的浓度较为敏感，所以 Mg^{2+} 的浓度对PCR产物的特异性和产量影响明显。过量的 Mg^{2+} 会导致酶催化非特异产物的扩增，而 Mg^{2+} 浓度过低，又会使 *Taq* 酶的催化活性降低。一般 Mg^{2+} 浓度在0.5～2.5mmol/L之间。如果溶液中存在EDTA，或在引物贮备液中有其他螯合物或者DNA模板，会干扰 Mg^{2+} 的浓度。这时就要适当提高 Mg^{2+} 的用量。

（六）dNTP浓度

dNTP的质量与浓度和PCR扩增效率有密切关系，dNTP呈颗粒状，如保存不当易降解，失去生物学活性。dNTP溶液呈酸性，使用时应配成高浓度贮存液，并以1mol/L的

NaOH 或 1mol/L 的 Tris-HCl 缓冲液将其 pH 调节到 7.0～7.5，最初的贮存液可稀释到 10mol/L，小量分装后于－20℃冰冻保存。多次冻融会使 dNTP 降解。PCR 反应体系中每种 dNTP 的终浓度为 50～200μmol/L，在此范围内，扩增产物量、特异性与合成忠实性之间的平衡最佳。dNTP 浓度过高会促进错误掺入，过低又影响产量。50～200μmol/L 的 dNTP 前体足够合成 5μg DNA。4 种 dNTP 的终浓度应该相同，任何一种的浓度明显偏高或偏低时，就会导致链延伸时的错误掺入增加，过早终止合成反应。如在 100μl 的反应体系中，4 种 dNTP 的浓度为 20μmol/L 可基本满足合成 2.6μg DNA 或 10pmol/L 的 400bp 序列的需要。另外，dNTP 的类似物也可加入 PCR 反应体系中，如 5-溴化脱氧尿嘧啶、生物素化脱氧尿嘧啶核苷或地高辛化脱氧尿嘧啶核苷，与 dNTP 的比例以 1∶3 加入，生成的 PCR 产物即为非放射性标记的核酸探针，用于核酸探针杂交实验。

（七）其他辅助试剂

PCR 反应中加入一定浓度的添加剂如 DMSO（二甲基亚砜）、甘油或甲酰胺等，可提高 PCR 扩增效率及特异性。这些试剂对 PCR 的具体影响机制尚不清楚，推测可能是其消除了引物和模板的二级结构，降低 DNA 双链的解链温度，使 DNA 双链变性完全所致。同时，添加剂还可增进 DNA 复性时的特异配对，增加或改变 DNA 聚合酶的稳定性等，提高 PCR 扩增效率。另外，反应中加入牛血清白蛋白（100μg/ml）、明胶（0.01%）或 Tween 20（0.05%～0.1%）有助于酶的稳定，反应中加入 5mmol/L 的二硫苏糖醇（DTT）也有类似作用，这对扩增长片段（此时延伸时间长）的 PCR 反应尤其有利。

对于不同的 PCR 反应体系，添加剂的浓度及其对 PCR 扩增的影响是不同的，当添加剂的浓度超过某一范围时，反而会抑制 PCR 扩增。表 3-1 归纳了几种添加剂对 PCR 反应的影响。

表 3-1　影响 PCR 反应的添加剂的浓度

名称	产生抑制作用的浓度/%	具有促进作用的浓度/%	名称	产生抑制作用的浓度/%	具有促进作用的浓度/%
DMSO	>10	5	甘油	>20	10～15
PEG	>20	5～15	Tween 20	未测定	0.1～2.5
甲酰胺	>10	5			

五、扩增产物的检测分析

目前检测 PCR 扩增产物的方法包括凝胶电泳、核酸探针杂交、酶切图谱分析、单链构型多态性分析、核酸序列分析等，其中凝胶电泳是检测 PCR 扩增产物最常用和最简便的方法之一。凝胶电泳主要有琼脂糖凝胶电泳和聚丙烯酰胺凝胶电泳两种常用技术。

（一）琼脂糖凝胶电泳

琼脂糖凝胶电泳是检测 PCR 扩增产物最常用的方法之一。凝胶根据需要可配制成各种浓度，进行不同目的的电泳。不同浓度的琼脂糖凝胶，可以分离不同大小的 DNA 片段（详见本书相关章节）。

核酸凝胶电泳结果的检测方法有溴化乙锭染色、银染色及同位素放射自显影等，其中溴化乙锭染色法较为普遍。溴化乙锭（ethidium bromide，简称 EB）是一种荧光剂，由于溴化乙锭分子插入在 DNA 双螺旋结构的两个碱基之间后，能形成一种在紫外光激发下发出很强橙红色荧光的络合物，所以十分容易观察。检测的灵敏度非常高，1μg/ml 的溴化乙锭溶液

可检出 10ng 或更少的 DNA 样品。溴化乙锭产生的荧光在紫外光源下放置时间过长能被猝灭，也容易受一些化学物质的污染而猝灭。溴化乙锭是一种 DNA 诱变剂，使用时应注意避免与皮肤接触。实验室中的 EB 污染物应妥善处理。EB 溶液的污染物不能直接倒入下水道及垃圾中，含有 EB 的凝胶应在干燥后烧毁。

用琼脂糖凝胶电泳法测定 DNA 片段的分子量，是在同一块凝胶板的样品槽中加待测样品，另一槽中加一个标准分子量样品，同时进行电泳，然后用溴化乙锭染色，在紫外灯下比较样品与标准品的位置，即可估计出待测样品的分子量大小范围。实验室常用的 DNA 分子量标准品有 pBR322 和 λDNA 的酶切片段。

（二）聚丙烯酰胺凝胶电泳

聚丙烯酰胺凝胶电泳具有琼脂糖凝胶电泳所不具备的优点：①分辨率很高，长度仅相差 0.2%（即 500bp 中的 1bp）的 DNA 分子即可分开；②能装载的 DNA 量远大于琼脂糖凝胶，一个加样槽（1cm×1mm）中加入 10μg DNA 也不会明显影响分辨率；③从聚丙烯酰胺凝胶中回收的 DNA 纯度很高。另外，可以采用银染色法染聚丙烯酰胺凝胶中的 DNA 和 RNA，其灵敏度比溴化乙锭染色法高 2～5 倍，而且避免溴化乙锭迅速退色的弱点，银染的凝胶干燥后可长期保存。PCR 扩增指纹图、多重 PCR 扩增、PCR 产物的酶切分析及利用产物的限制性片段长度多态性（PCR-RFLP）进行基因分型等时，聚丙烯酰胺凝胶电泳是一种理想的手段。不同浓度的聚丙烯酰胺凝胶，可以分离 DNA 片段大小的范围参数列于表 3-2。

表 3-2　聚丙烯酰胺凝胶浓度与分离 DNA 片段大小的关系

丙烯酰胺浓度/%	有效分离范围/bp	溴酚蓝位置相当于双链 DNA 片段/bp	丙烯酰胺浓度/%	有效分离范围/bp	溴酚蓝位置相当于双链 DNA 片段/bp
3.5	1000～2000	100	12.0	40～200	20
5.0	80～500	65	15.0	25～150	15
8.0	60～400	45	20.0	6～100	12

（三）核酸探针杂交鉴定法

核酸探针杂交法是鉴定 PCR 扩增产物特异性的一种重要手段。当扩增出现多条带或存在非特异性扩增掩盖作用时，尤其需要核酸探针辅助鉴别。从杂交类型看，点杂交除可鉴定扩增产物外，还有助于产物的分型及法医学鉴定；Southern 印迹杂交可以从非特异扩增产物中鉴定出特异性产物的大小，增加 PCR 检测的特异性和灵敏度。

（四）限制性内切酶分析法

酶切分析是鉴别 PCR 扩增产物特异性的一种简便方法。根据目标基因的已知序列资料可以查出所包含的酶切位点，用某种限制性内切酶消化扩增产物后进行电泳，观察消化片段的数目及大小是否与序列资料相符，从而确定产物的特异性。酶切分析的另一种用途是遗传病的诊断和传染病病原体的基因分型。酶切位点的改变是序列差异的遗传路标，因此可利用高频率切点的限制酶消化 PCR 扩增产物，根据限制性片段长度多态性（RFLP）进行目标基因分析和分型。

特异性鉴别可选用识别 6 个碱基的限制性内切酶，从已知序列资料中查出。基因分型可用识别 4 个碱基的限制性内切酶，常用的有 *Alu* Ⅰ、*Msp* Ⅰ、*Taq* Ⅰ、*Hinf* Ⅰ、*Mbo* Ⅰ、*Dde* Ⅰ、*Mse* Ⅰ、*Bbv* Ⅰ、*Mae* Ⅰ、*Mae* Ⅲ、*Fnu*4H Ⅰ、*Hha* Ⅰ、*Rsa* Ⅰ等。

（五）单链构型多态性分析法

单链构型多态性分析法（PCR-single strand conformation polymorphism，简称 PCR-SSCP）是一种简便快速的 PCR 扩增产物分析方法。其基本原理是双链 DNA 的 PCR 扩增产物，经变性处理成单链 DNA，加样于非变性聚丙烯酰胺凝胶进行电泳，由于 DNA 分子在凝胶电泳中的迁移率与其分子量和空间构型有关，在非变性条件下，单链 DNA 因分子内力作用形成卷曲构型，这种二级结构与单链 DNA 的序列有关，因此，单链 DNA 的电泳迁移率受到 PCR 产物二级结构的影响。电泳结束后，单链 DNA 带在凝胶中位置的差异反映了 PCR 产物序列的差异。

六、PCR 条件优化

PCR 扩增大量目的产物是理想的结果，但在未找到最佳扩增条件时可能产生许多不需要的非特异性产物或没有扩增到任何产物。这就需要对 PCR 条件进行优化，以达到预期的扩增目的。

（一）常规的优化方法

适当地改变对引物-模板忠实性或对引物延伸有影响的因素。

1. 有利于增加特异性的条件

如降低 Mg^{2+} 浓度，降低 dNTP 浓度，优化 pH，降低 *Taq* 酶用量，减少循环中各部分的时间或循环数，增加退火温度，减少或去除抑制剂等。

2. 加入或优化增强剂

加入一系列适当浓度的辅助物，如 DMSO（二甲基亚砜）、PEG6000、甘油、非离子去污剂、甲酰胺和牛血清白蛋白到反应体系中可以提高特异性和产量。

3. 优化引物设计

借助于计算机引物设计软件来达到引物的优化。

4. 热启动技术

在 PCR 反应的第一个循环中待温度升高超过模板 T_m 值（80℃）后才加一、两种关键试剂。例如在覆盖有矿物油的小试管的反应体系中先不加入 *Taq* DNA 聚合酶，而到第一个循环的变性阶段温度超过 80℃ 以后再加，这样，可以大大减少引物二聚体和非特异性配对的形成，提高 PCR 的特异性。还可利用石蜡将需要后加的成分与反应液隔开，当加热至石蜡被融化后，上面的后加成分即进入反应体系；还有，利用加入抗 *Taq* DNA 聚合酶的单克隆抗体于 PCR 反应管中，在温度升高到将中和抗体变性失活之前，抗体阻遏 *Taq* 酶活性，防止反应开始。

（二）降落 PCR

降落 PCR（touchdown PCR，TD-PCR）是连续改变反应管退火温度，以找到最佳扩增条件的方法。开始的退火温度选择为高于估计的 T_m 值，随着循环的进行，退火温度逐渐降到 T_m 值，并最终低于这个水平。TD-PCR 有利于确保第一个引物-模板杂交过程发生在最为互补的、产生目的扩增产物的反应物之间。当目的扩增产物已开始指数扩增时，即使剩下循环的退火温度最终降到非特异杂交的 T_m 值，此时目的产物的扩增已经处于非常优势的地位。一般在编设 TD-PCR 程序时，退火温度的范围应该超过 15℃，从高于 T_m 值几度降到低于它 10℃左右。

七、PCR 常见问题与对策

（一）PCR 污染

PCR 技术的敏感性极高，极微量的污染即可导致假阳性的产生，因此了解 PCR 污染的原因，采取相应措施预防和消除 PCR 污染就十分重要。

PCR 污染主要是随机污染，包括 PCR 反应前污染及反应后污染。

1. PCR 反应前污染

主要是样品 DNA 的交叉污染。提取 DNA 使用的仪器上残留的杂质 DNA 或反应管中残留的 PCR 产物、PCR 操作者皮肤或头发等均可引起样品污染。其中 PCR 产物的污染，也称残留污染（carryover），是引起样品交叉污染的主要因素。明胶、*Taq* DNA 聚合酶等 PCR 反应试剂等均有可能带有杂质 DNA，因此也可导致 PCR 反应前污染。

2. PCR 反应后污染

PCR 反应后污染主要来自于 PCR 产物检测过程，如电泳上样、点杂交点样时加样器吸头之间的污染等。

3. 潜在的 PCR 污染源

主要包括 3 方面：①仪器设备的污染，样品收集器、微量加样器、点杂交器、离心管、离心机、切片机、pH 计、水浴锅、高压锅及 PCR 仪等的污染；②试剂污染，乙醇、液氮、氯仿、酶及其他生物制品（如明胶等）的污染；③操作者的污染，操作者的皮肤、头发等均可引起 PCR 污染。

（二）污染的预防

进行 PCR 反应时，遵照下列规则有助于防止 PCR 的污染。

① PCR 的前处理和后处理在不同的房间或不同的隔离工作台上进行。即整个 PCR 操作要在不同的隔离区（标本处理区，PCR 扩增区，产物分析区）进行，特别是阳性对照需在另一个隔离环境中贮存、加入。

② 试剂要分装，每次取完试剂后盖紧塞子。PCR 反应时将反应成分制备成混合液，然后再分装到不同的反应管中，这样可减少操作，避免污染。

③ 改进实验操作，戴一次性手套，PCR 实验应配置专用的微量可调加样器，这套加样器绝不能用于 PCR 反应产物的分析。用一次性移液器吸头、反应管，避免反应液的飞溅等。

④ 检查结果的重复性。

⑤ 经常处理仪器设备等潜在的 PCR 污染源，减少污染可能。

（三）实验中的对照

每一次试验都需要设置严格的对照。

阳性对照：阳性模板。

阴性对照：阴性模板或不加模板。

（四）扩增反应总是阴性结果（无产物）时应采取的措施

① 取 10 μl 扩增混合液作模板再进行 PCR 扩增。

② 增加 *Taq* DNA 聚合酶的浓度。

③ 增加靶 DNA 量。

④ 若模板为粗制品，提纯样品。

⑤ 增加扩增循环次数。

⑥ 适当降低退火温度。

（五）出现非特异性产物时应采取的措施

① 增加退火温度。

② 减少 *Taq* DNA 聚合酶的浓度。

③ 减少退火及延伸时间。

④ 减少引物浓度。

⑤ 减少扩增循环次数。

第二节　PCR 技术的扩展

随着 PCR 技术的发展，通过对 PCR 技术的大量改进派生出了一些新的相关技术和改良方法，开发了 PCR 技术的许多新用途。

一、逆转录 PCR

RT-PCR（reverse transcriptase PCR，RT-PCR）是一种 RNA 逆转录和 PCR 结合起来建立的 RNA 的聚合酶链反应。RT-PCR 使 RNA 检测的敏感性提高了几个数量级［较 Northern 印迹杂交敏感（3～6）×10^3 倍］，也使一些极微量 RNA 样品分析成为可能。

RT-PCR 的关键步骤是 RNA 的逆转录，要求 RNA 模板必须是完整的，不含 DNA、蛋白质等杂质。用于该反应的引物可以是随机六聚核苷酸或寡聚脱氧胸苷酸（oligo dT），还可以是针对目的基因设计的特异性引物（GSP）。研究提示，使用随机六聚引物延伸法的结果较为恒定，并能引起靶序列的最大扩增。一般而言，1μg 细胞 RNA 足以用于扩增稀有 mRNA 序列（1～10 拷贝/细胞）。一个典型的哺乳动细胞含约 10pg RNA。1μg RNA 相当于 10 万个细胞的 RNA 总量。

1. 逆转录反应

① 20μl 反应体系包括：1mmol/L dNTP；100mmol/L Tris-HCl，pH8.4；50mmol/L KCl，2.5mmol/L $MgCl_2$；100mg/ml BSA；100pmol 随机六聚寡核苷酸或 oligo dT 引物；1～9μl RNA 样品，80～200U 逆转录酶（M-MLV 或 AMV）。

② 混匀后，置室温 10min，然后置 42℃，30～60min。

③ 若需重复一次逆转录反应，则 95℃ 3min，然后补加 50U 逆转录酶后重复第②步。

④ 95℃ 10min 灭活逆转录酶，并使 RNA-cDNA 杂交体解链。

2. PCR 操作

① 将 20μl 逆转录物加入 80μl 含上、下游引物各 10～50pmol 的 PCR 缓冲液［100mmol/L Tris-HCl（pH8.4）；50mmol/L KCl；2.5mmol/L $MgCl_2$；100mg/ml BSA］中混匀。

② 加 1～2U *Taq* DNA 聚合酶和 100μl 石蜡油。

③ 根据待分析 RNA 的丰度决定循环数的多少。

④ 用 200～300μl TE 缓冲液饱和的氯仿抽提去除石蜡油。

⑤ 离心，取上层水相 5～10μl，用凝胶电泳分析 PCR 产物。

RT-PCR 也可以在一个系统中进行，称为一步法扩增（one step amplification），它能检测低丰度 mRNA 的表达。即利用同一种缓冲液，在同一体系中加入逆转录酶、引物、*Taq* 酶、4 种 dNTP，直接进行 mRNA 逆转录与 PCR 扩增。由于发现 *Taq* 酶不但具有 DNA 聚合酶的作用，而且具有逆转录酶活性，因此在同一体系中直接使用 *Taq* 酶，以 mRNA 为模板进行逆转录和其后的 PCR 扩增，使 mRNA-PCR 步骤更为简化，所需样品量减少到最低限度，对临床少量样品的检测非常有利。用一步法扩增可检测出总 RNA 中小于 1ng 的低丰度 mRNA。该法还可用于低丰度 mRNA 的 cDNA 文库的构建及特异 cDNA 的克隆，并有可能与 *Taq* 酶的测序技术相结合，使得自动逆转录、基因扩增与基因转录产物的测序在一个试管中进行。

二、定量 PCR

随着 PCR 技术的发展与广泛应用，利用 PCR 的定量技术也取得了长足的发展。早期主要采用外参照的定量 PCR 方法，这种方法因影响因素较多，现已逐渐被内参照物定量方法和竞争 PCR 定量方法所取代，并在此基础上发展了荧光定量 PCR 方法，使 PCR 定量技术提升到新的水平。

（一）利用参照物的定量方法

参照物是在定量 PCR 过程中使用的一种含量已知的标准品模板。按其性质的不同可分为内参照和外参照。外参照是序列与检测样品相同的标准品模板，外参照物与待检样本的扩增分别在不同的管内进行。通过一系列不同稀释度的已知含量外参照的扩增，建立外参照扩增前含量与扩增产物含量之间的标准曲线，以此用于未知样品的定量。因此，外参照定量 PCR 属非竞争性定量 PCR。

利用内参照物的定量方法与上述方法不同，它是将参照物与待检样本加于同一反应管中，参照物作为标准品模板与待检样品共用或不共用同一对引物，进行 PCR 扩增。按照内参照在扩增中是否与待检样品共用同一对引物，两个模板的扩增是否存在竞争性，内参照又可分为竞争性内参照和非竞争性内参照。

若内参照与待检样本共用同一对引物，两模板的扩增存在竞争性，则称为竞争性内参照，在这种条件下进行的定量 PCR 称竞争性定量 PCR。理想的竞争性内参照应具备以下特点：①竞争性内参照模板与检测样品能共用同一对引物；②内参照与待检样品长度相近，扩增效率相同；③内参照扩增产物能方便地与待检样本扩增产物分开。在竞争性 PCR 中，首先通过一系列不同稀释度的已知含量检测样品与已知含量的内参照混合共扩增，做出扩增前检测样品含量和内参照含量的比例与检测样品扩增产物含量的相关曲线，以此作为标准曲线用于待检样本的定量。若内参照与待检样本不共用同一对引物，两模板的扩增不存在竞争性，这时内参照又称为非竞争性内参照。这种 PCR 因其不存在竞争性，因而属于非竞争性 PCR。

（二）竞争性定量 PCR

采用内参照竞争性定量方法，在准确性方面优于外参照的定量方法，是目前较理想的一种定量方法。其产物分析可采用探针杂交法，或采用电泳法、HPLC 和荧光法等。由于采用内参照的竞争性定量，克服了常规 PCR 扩增效率不稳定的缺陷，各管间差异得以避免。

竞争性 PCR 方法的总策略是采用相同的引物，同时扩增靶 DNA 和已知浓度的竞争模板，竞争模板与靶 DNA 大致相同，但其内切酶位点或部分序列不同，用限制性内切酶消化

PCR 产物或用不同的探针进行杂交即可区分竞争模板与靶 DNA 的 PCR 产物，因竞争模板的起始浓度是已知的，通过测定竞争模板与靶 DNA 二者 PCR 产物便可对靶 DNA 进行定量。

利用竞争 PCR 亦可进行 mRNA 的定量。先以 mRNA 为模板合成 cDNA，再用竞争 PCR 对 cDNA 定量。但当逆转录效率低于 100%时，通过测定样品中 cDNA 进行 mRNA 定量，则测定结果会偏低。

（三）荧光定量 PCR

荧光定量 PCR（FQ-PCR）是新近出现的一种定量 PCR 检测方法，可采用外参照的定量方法，也可采用内参照的定量方法。荧光定量 PCR 采用各种方法用荧光物质标记探针，通过 PCR 显示扩增产物的量。由于荧光探针标记方法的不同，可分为不同的方法，并可借助专用的仪器实时监测荧光强度的变化。因此，这种方法可以免除标本和产物的污染，且无复杂的产物后续处理过程，因而该方法准确、高效、快速。

早期荧光定量 PCR 有许多局限性，如不能准确定量或由于太灵敏而容易交叉污染，产生假阳性。直到最近，荧光能量传递技术（fluorescence resonance energy transfer，FRET）应用于 PCR 定量后，上述问题才得到较好地解决。FRET 是指通过供、受体发色团之间偶极-偶极相互作用，能量从供体发色团转移至受体发色团，使受体发光。以下介绍两种常用的测定方法。

1. TaqMan 技术

TaqMan 是由 PE 公司开发的荧光定量 PCR 检测技术，它在普通 PCR 原有的一对引物基础上，增加了一条特异性的荧光双标记探针（TaqMan 探针）。TaqMan 技术的工作原理：PCR 反应系统中加入的荧光双标记探针，可与两引物包含序列内的 DNA 模板发生特异性杂交，探针的 5′端标以荧光发射基团 FAM（6-羧基荧光素，荧光发射峰值在 518nm 处），靠近 3′端标以荧光猝灭基团 TAMRA（6-羧基四甲基罗丹明，荧光发射峰值在 582nm 处），探针的 3′末端被磷酸化，以防止探针在 PCR 扩增过程中被延伸。当探针保持完整时，猝灭基团抑制发射基团的荧光发射。发射基团一旦与猝灭基团发生分离，抑制作用即被解除，518nm 处的光密度增加而被荧光探测系统检测到。复性期探针与模板 DNA 发生杂交，延伸期 *Taq* 酶随引物延伸沿 DNA 模板移动，当移动到探针结合的位置时，发挥其 5′→3′外切酶活性，将探针切断，猝灭作用被解除，荧光信号释放出来（图 3-4）。荧光信号即被特殊的仪器接收。PCR 进行一个循环，合成了多少条新链，就水解了多少条探针，释放了相应数目的荧光基团，荧光信号的强度与 PCR 反应产物的量呈对应关系。随着 PCR 过程的进行，重复上述过程，PCR 产物呈指数形式增长，荧光信号也相应增长。如果每次

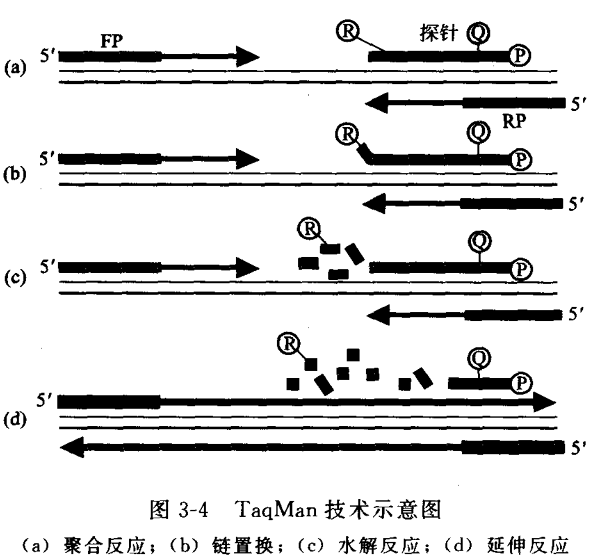


图 3-4　TaqMan 技术示意图

（a）聚合反应；（b）链置换；（c）水解反应；（d）延伸反应

（P：探针；R：FAM；Q：TAMRA；

FP：上游引物；RP：下游引物）

测定 PCR 循环结束时的荧光信号与 PCR 循环次数作图，可得一条“S”形曲线。如果标本中不含阳性模板，则 PCR 过程不进行，探针不被水解，不产生荧光信号，其扩增曲线为一水平线。

TaqMan 技术中对探针有特殊要求：①探针长度应在 20～40 个碱基，以保证结合的特异性；②GC 含量在 40%～60%，避免单核苷酸序列的重复；③避免与引物发生杂交或重叠；④探针与模板结合的稳定程度要大于引物与模板结合的稳定程度，因此探针的 T_m 值要比引物的 T_m 值至少高出 5℃。

TaqMan 技术在基因表达分析、血清病毒定量分析、人端粒酶 mRNA 定量分析及基因遗传突变分析等领域有广泛应用。但也存在缺陷，主要包括：①采用荧光猝灭及双末端标记技术，因此猝灭难以彻底，本底较高；②利用酶外切活性，因此定量时受酶性能影响；③探针标记成本较高，不便普及应用。

2. 分子信标技术

分子信标技术（molecular beacon）也是在同一探针的两末端分别标记荧光分子和猝灭分子。通常，分子信标探针长约 25 nt（nucleotide，核苷酸），空间上呈茎环结构。与 TaqMan 探针不同的是分子信标探针的 5′和 3′末端自身形成一个 8 个碱基左右的发夹结构，此时荧光分子和猝灭分子邻近，因此不会产生荧光。当溶液中有特异模板时，该探针与模板杂交，从而使探针的发夹结构打开，于是溶液便产生荧光（图 3-5）。荧光的强度与溶液中模板的量成正比，因此可用于 PCR 定量分析。该方法的特点是采用非荧光染料作为猝灭分子，因此荧光本底低。

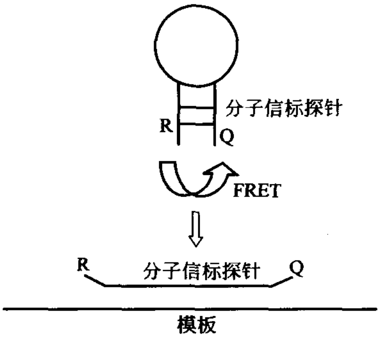


图 3-5　分子信标技术示意图

常用的荧光猝灭分子对是 5′-(2′-氨乙基氨基萘-1-磺酸)(EDANS) 和 4′-(4′-二甲基氨基叠氮苯) 苯甲酸 (DAB-CYL)，采用 DAB-CYL 作猝灭剂是由于它对多种荧光素都有很强的猝灭效率。当受到 336nm 紫外光激发时，EDANS 发出波长为 490nm 的亮蓝荧光；DAB-CYL 是非荧光分子，但其吸收光谱与 EDANS 的发射荧光光谱重叠。只有分子信标时，EDANS 和 DAB-CYL 距离非常接近，足以发生 FRET，EDANS 受激发产生的荧光转移给 DAB-CYL 并以热的形式散发，不能检测到荧光；相反，当有靶核酸存在时才可检测到荧光。

分子信标技术不足之处是：①杂交时探针不能完全与模板结合，因此稳定性差；②探针合成时标记较复杂。分子信标技术结合不同荧光标记，可用于基因多突变位点的同时分析。

三、重组 PCR

重组 PCR（recombinant PCR，R-PCR）可用 PCR 法在 DNA 片段上进行定点突变，突变的产物（即扩增物）中含有与模板核苷酸序列相异的碱基，用 PCR 介导产生核苷酸的突变包括碱基替代、缺失或插入等。

重组 PCR 操作需要 2 对引物：“左方”PCR 的一对引物为 a 和 b′，b′中含有一个“突变碱基”；“右方”PCR 一对引物 b 和 c，b 中含有一个和引物 b′中的“突变碱基”相互补的碱基。先用 2 对引物分别对模板进行扩增。除去引物后将 2 种扩增产物混合，变性并复性后进

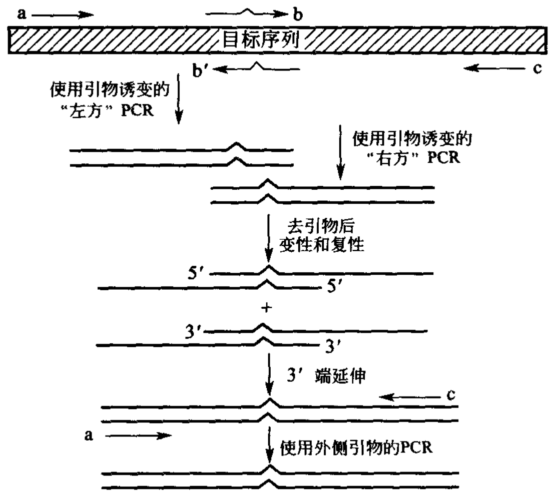


图 3-6 重组 PCR 造成特定碱基置换示意图

行延伸，然后再加入外侧引物 a 和 c，经常规的 PCR 循环后，便能得到中间部位发生特定突变的 DNA 片段（图 3-6）。重组 PCR 造成 DNA 片段的插入或缺失与其造成特定碱基的置换在操作上相类似。重组 PCR 可制备克隆突变体，用作研究基因的功能。

四、反向 PCR

常规 PCR 是扩增两引物之间的 DNA 片段，反向 PCR（reverse PCR）是用引物来扩增两引物以外的 DNA 片段。一般，先用限制性内切酶酶解 DNA（目的基因中不存在该酶的酶切位点，且片段应短于 2～3kb），然后用连接酶使带有黏性末端的靶片段自身环化，最后用一对反向引物进行 PCR，得到的线性 DNA 将含有两引物外侧的未知序列，如图 3-7 所示。该技术可对未知序列扩增后进行分析，如探索邻接已知 DNA 片段的序列。反向 PCR 已成功地用于仅知部分序列的全长 cDNA 克隆，Picter 曾用此法扩增基因文库的插入 DNA。

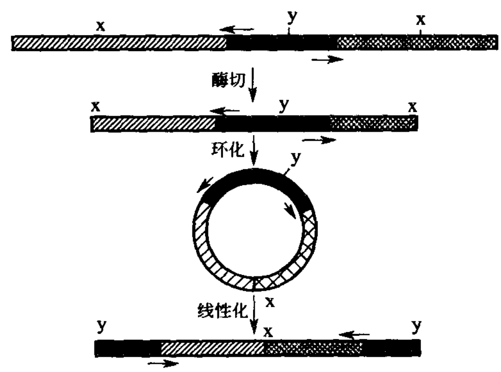


图 3-7 反向 PCR 示意图
（x、y 为两种限制性内切酶）

五、不对称 PCR

不对称 PCR 的基本原理是采用不等量的一对引物，经若干次循环后，低浓度的引物被消耗尽，以后的循环只产生高浓度引物的延伸产物，结果产生大量的单链 DNA（ssDNA）。这两种引物分别称为限制性引物和非限制性引物，其最佳比例一般是 1/50～1/100，因为 PCR 反应中使用的两种引物浓度不同，因此称为不对称 PCR。

不对称 PCR 主要为测序制备 ssDNA，其优点是不必在测序之前除去剩余引物，因为量很少的限制性引物已经耗尽。多数学者认为，用 cDNA 经不对称 PCR 进行 DNA 序列分析是研究真核 DNA 外显子的好方法。

六、复合 PCR

该技术就是在一个反应体系中同时加入多对套叠的 PCR 引物，从而获得最大效率的 PCR 产物。最新的复合 PCR 反应记录是由 Chamberlain 创造的，他证明 PCR 可以在 9 个不同位点上同时开始反应，采用 9 套不同的引物从而得到 DMD 基因片段。多数学者认为，要想获得高纯度特异性 PCR 产物，最好用两对相互套叠的引物，这样得到的 PCR 扩增 DNA 比一对引物扩增的 DNA 特异性强。

七、着色互补PCR

应用PCR技术扩增感兴趣的DNA区带时，通常需要对PCR扩增产物作进一步的分析、鉴定，如电泳、序列分析和DNA分子杂交。在多重PCR试验中，一次试验还不容易区分长短相近的多种基因成分，着色互补（color complementation assay）或称荧光PCR（fluroscent PCR）可用于解决这一问题。

荧光PCR的原理是用不同的荧光染料，如红色的罗丹明（6-carboxy-rhodamin，POX）和绿色荧光素（5′-carboxyfluorescein，FAM）等分别标记于不同的寡核苷酸引物上，同时扩增多个DNA区段。反应完毕后，利用分子筛除去延伸的引物。用紫外线照射扩增产物，就能显示某一种或几种荧光染料颜色的组合，如果某一DNA区带缺失，则会缺乏相应的颜色，据此可以很快诊断是否某种基因缺失，或是发现某些感染的病毒等。该技术为PCR成为临床自动化诊断方法打下基础。

八、锚定PCR

当待扩增的DNA或RNA序列本身或旁侧序列不清楚时，可以采用锚定PCR技术对目的片段进行扩增。具体方法是：分离细胞总RNA或mRNA，在逆转录酶作用下合成cDNA，通过DNA末端转移酶在cDNA 3′末端加上poly(dG)尾。与此poly(dG)相对应的锚定引物为poly(dC)，为保证扩增特异性，选择此段poly(dC)在12聚体以上为佳，5′端可带上某些限制酶序列或其他序列信息。在与基因特异配对的引物参与下，可以扩增此带同源多聚物尾巴的cDNA序列。

九、原位PCR

原位PCR（in situ PCR）是一种把原位杂交的细胞定位技术与PCR的高灵敏度性相结合的技术。该技术可以获得靶基因（靶DNA或RNA）在细胞的特异定位信息，而靶基因的细胞定位在多数情况都是重要的信息，例如了解HIV-DNA在血清阳性无症状病人哪些细胞型是病毒的原发灶等方面十分重要。

简要步骤：组织切片和细胞样品制备→蛋白酶消化→DNA酶消化（用于原位PCR中的逆转录）→逆转录步骤→PCR→检测。

十、膜结合PCR

膜结合PCR（membrane-bound PCR）是将模板DNA经一定处理后固定到尼龙膜或硝酸纤维素膜上，将膜上固定的DNA直接用于PCR扩增，此法扩增效率较低。因此，较液相PCR需多出10～20个循环。

膜结合PCR特别适用于模板量极微少，或DNA严重污染的情况（由于DNA结合在膜上，可通过彻底漂洗来纯化模板）。另外，若PCR本底较高，可经电泳使待分析片段吸印于膜上，进行固相PCR。

十一、固着PCR

固着PCR（sanchored PCR，SA-PCR）是利用固相化的引物进行PCR，使扩增的产物也固相化，这有利于产物的分离。含Aminolik Ⅱ的寡核苷酸引物能与固相支持物琼脂糖珠

相连。作为分离 RNA 的亲和基质，逆转录后，cDNA 便间接地结合到琼脂糖珠上，相当于构建了一个特异“cDNA 文库”。另外，固相化引物也易于在不同缓冲系统中操作。

十二、增效 PCR

增效 PCR（booster PCR）是一种双相扩增法，能有效地阻止引物二聚体的形成，这样不但节约了引物和酶，还使靶序列产量得到了提高。增效 PCR 方法的原理：在适当稀释的引物条件下，反应中引物相对碰撞的机会减少，而使引物和模板复性。但引物的减少会明显影响待扩序列的产量，所以在扩增一定循环数后补加引物，再于正常 PCR 条件下扩增，使产物呈指数增加，操作时分两步：第一步是在引物和模板均少的状态下进行，使引物较模板的摩尔浓度高 107 倍，进行 15～20 个循环；第二步补加引物至 0.1μmol/L，再于常规条件下进行 30～50 个循环。由此可见，增效 PCR 的第一步主要目的是增加模板量，有效防止第二步扩增时加入的过多引物间的相互作用，阻止了引物二聚体的形成，增加了特异性；第二步则可大大提高特异性靶序列的产量。

第三节　PCR 技术的应用

一、PCR 技术在分子生物学中的应用

1. 基因克隆

基因的克隆和分离是分子生物学和细胞生物学研究中必不可少的手段。运用 PCR 技术进行基因克隆和亚克隆比传统的方法具有更大的优点。由于 PCR 可以对单拷贝的基因放大上百万倍，产生微克（μg）级的特异 DNA 片段，从而可省略从基因组 DNA 中克隆某一特定基因片段所需要的 DNA 的酶切、连接到载体 DNA 上、转化、建立 DNA 文库及基因的筛选、鉴定、亚克隆等烦琐的实验步骤。

只要知道目的基因的两侧序列，通过一对和模板 DNA 互补的引物，可以有效地从基因组 DNA 中、mRNA 序列中或已克隆到某一载体上的基因片段中扩增出所需的 DNA 序列。为了克隆操作的方便或保证克隆后基因的方向正确性，可以在设计引物时在其 5′端作一些修改，或加上一段不完全与模板互补的序列。这个添加的序列可以是带限制性内切酶切点的连接子，或某种启动子序列、起始或终止密码。经过 PCR 扩增后，添加的序列最终可整合到新合成的产物中，便于下一步的重组克隆以及基因的表达和调控研究。如果引入的序列是包含限制性内切酶位点的连接子，PCR 产物经相应的内切酶消化后，即可进行定向克隆。

与传统的 DNA 克隆方法比较，采用 PCR 的 DNA 克隆方法省时、省力，但它需要知道目的基因两侧序列的信息。

2. 重组 PCR

在分子生物学研究中常常需要将两个不同的基因融合在一起。通过 PCR 反应可以比较容易地实现这一目的。相关内容请见上一节中的“重组 PCR”部分。

3. DNA 序列的测定

目前广泛采用的 DNA 测序方法有化学法和双脱氧法两种，它们对模板的需要量比较大。传统的模板制备方法是将含目的基因的 DNA 片段进行酶切，建立基因库，筛选克隆，并亚克隆到 M13 噬菌体载体上，经噬菌体产生单链 DNA。这是一项比较复杂的工作。利用 PCR 方法可以比较容易地测定位于两个引物之间的序列。利用 PCR 测序有下列几种方法。

（1）双链直接测序　将 PCR 产物进行电泳回收或利用分子筛等方法除去小分子物质。采用热变性或碱变性的方法使双链 DNA 变成单链，与某一方向的引物退火，用 Sanger 的双脱氧法即可测出一定长度的 DNA 序列。这种方法测序时，一次测定的距离较短，同时容易出现假带。

（2）双链克隆后测序　为避免双链直接测序的缺点，可以将 PCR 扩增产物直接克隆到 M_{13} 载体中，通过提取单链后进行序列测定。

（3）基因扩增转录测序　在 PCR 时，使用一个 5′端带有 T_7 启动子的序列，扩增后在 T_7 RNA 酶作用下合成 mRNA 作为模板，用逆转录酶测序。

（4）不对称 PCR 产生单链测序　在 PCR 反应中加入不等量的一对引物，经过若干循环后，其中一种引物被消耗尽，在随后的循环里，只有一种引物参与反应，结果合成了单链。这是一种简单又可靠的测序方法，此反应不必在测序之前除去剩余的引物，因为量少的引物在反应中已经消耗掉，也不用担心测出由于 *Taq* DNA 聚合酶碱基识别能力限制所引起的基因突变，因为某个随机突变在所有单链中所占比例总是很少的。

4. 用于基因定量

应用 PCR 技术可以定量监测标本中靶基因的拷贝数，这在研究基因的扩增等方面具有重要意义，这种方法称为差示 PCR（differential PCR）。它是将目的基因和一个单拷贝的参照基因置于同一个试管中进行 PCR 扩增。电泳分离后呈两条区带，比较两条区带的丰度；或在引物 5′端标记上放射性核素后，通过检测两条区带放射性强度即可测出目的基因的拷贝数。利用差示 PCR 还可以对模板 DNA（或 RNA）的含量进行测定，在一系列 PCR 反应中，分别加入待测模板 DNA 和参照 DNA 片段。参照 DNA 片段基本上按待测 DNA 的方式进行构建，但增加了一小段内部连接顺序，这样使得两种 DNA 片段的 PCR 产物可在凝胶电泳上分开。由于这两种不同的 DNA 片段可以在同一种寡核苷酸引物作用下同步扩增，因此可以避免因使用不同的引物所引起的可能误差。这两种扩增产物的相对量就反应了在起始反应混合物中的目的 DNA 和参照 DNA 的相对浓度。

利用差示 PCR 同样可以做 mRNA 的定量。与 DNA 模板含量测定基本相似，所不同的是首先应提取总 RNA，加入一个与 mRNA 3′区互补的引物，在逆转录酶的作用下，合成 cDNA，其余的操作与 DNA 定量相同。利用 PCR 可以对 mRNA（如进行性肌萎缩蛋白 mRNA）作比较准确的定量，甚至连 Northern 印迹杂交未能发现的 mRNA 都可以得到检测。

二、PCR 技术在传染病病原体检测中的应用

传统的实验室诊断传染病病原体感染的方法主要有：①病原体培养法，需要时间较长，且对标本要求高，不能作早期诊断；②免疫学方法，在抗原量低时敏感度不够，有较强的交叉反应，且不能诊断潜伏感染；③核酸探针法，敏感性不足，在实用性方面颇受限制。

PCR是一种具有高敏感性、且特异性好的靶DNA快速检测方法，能对前病毒或潜伏期低复制的病原体特异性靶DNA片段进行扩增检测，只要标本中含有1fg靶DNA就可检出。可检出经核酸分子杂交呈阴性的许多标本，而且对标本要求不高，经简单处理后就能得到满意扩增。可做到早期及时诊断，对防止传染病流行有重要意义。

PCR可用于病原体的潜伏期、早期诊断；临床检测与监控；流行病学、分子传染病学研究；评价药物治疗效果。

PCR技术已在临床上成功地应用于多种病原体检测，如检测甲型肝炎病毒（HAV）、乙型肝炎病毒（HBV）、丙型肝炎病毒（HCV）、戊型肝炎病毒（HEV）、肾综合征出血热病毒（HFRSV）、人乳头瘤病毒（HPV）、单纯疱疹病毒（HSV）、人巨细胞病毒（HCMV）、人类免疫缺陷病毒（HIV）、人微小病毒B19、乙脑病毒、轮状病毒、风疹病毒（RV）、EB病毒、腺病毒（AV）、结核杆菌（TB）、嗜肺军团杆菌（aB）、淋球菌（NG）、沙门菌（SM）、脑膜炎球菌、霍乱弧菌、幽门螺旋杆菌（HP）、白色念球菌、梅毒螺旋体（TP）、沙眼衣原体（CT）、肺炎支原体（MP）、解脲支原体（UU）。

由于病原体检测常常是在大量宿主核酸存在的情况下进行的，因此具有病原体特异性的引物是扩增敏感而特异的关键。由于敏感性高的同时易因微量污染造成假阳性，所以应严格操作，同时每次设置阳性和阴性对照，以有效解决假阳性问题。

三、PCR在肿瘤相关基因检测中的应用

恶性肿瘤是严重危害人类健康的一种世界性的常见病和多发病，死亡率较高。致癌原因很多且很复杂，包括化学的、物理的及生物的因素，但其本质都是引起基因的变异而致癌。恶性肿瘤的发生、发展是由于组织细胞的遗传物质发生了改变，导致组织细胞的无限制生长，甚至转移到其他部位。这种遗传物质改变的原因分为两大类：一类是家族性遗传的，如抑癌基因杂合性丢失，常引起视网膜细胞瘤、结肠息肉瘤、多发性内分泌综合瘤、Wilms瘤、乳腺癌、肝癌等；另一类为后天所得，如某些因素引起遗传失调，发生控制细胞增殖的原癌基因的染色体重排，导致癌基因与抑癌基因之间的平衡被破坏而引发肿瘤。

细胞基因组的变化影响了控制细胞生长和分化的基因表达及功能，从而导致癌症，癌基因和抑癌基因统称为肿瘤相关基因或细胞调控基因。

癌基因的活化或抑癌基因的失活的原因包括点突变、缺失、片段的插入或易位等。可通过下列方法进行诊断检测。

1. 寡核苷酸探针杂交法

以点突变为中心设计出一条长19～20个碱基的标记的寡核苷酸链作探针，根据完全配对的碱基比有单个碱基错配的杂交稳定的原理，采用严格的杂交和洗涤条件，就可以确定待测DNA样本是否存在突变的碱基。该技术有高度特异性和敏感性，是分子生物学研究中常用的重要技术。

2. 限制性内切酶切片段长度多态性（RFLP）技术

由于基因的突变，导致限制性内切酶识别位点的消失、改变，而产生与正常等位基因不同长度的片段，据此可以判断基因的改变。

3. PCR-RFLP技术

以c-Ha-*ras*基因为例，引物设计在点突变位点上游区和下游区，经PCR扩增出100bp

长的片段，由于第12密码子的点突变使限制性内切酶*Hpa* Ⅰ的识别位点被破坏。若将扩增产物经*Hpa* Ⅰ消化后，健康人的DNA片段被切成为50bp，而发生突变基因的病人DNA的产物为100bp，经电泳就可以区别开来。

4. RNA酶A错配清除法

RNA酶A能识别单碱基错配的RNA-RNA或RNA-DNA双体。这时可用体外合成与正常序列互补的RNA探针。将探针和变性的肿瘤DNA或RNA杂交，然后用RNA酶A处理，可以在RNA-RNA或RNA-DNA错配部位切断RNA，使之形成2个片段，正常组织DNA经相同处理后就不会形成2个片段，据此可以互相区别开来。

5. 核酸序列分析法

目前大多数癌症都是在瘤组织块形成后才能被诊断。体液或活检组织的细胞学分析都可比较早地作出诊断，但所有的方法都依靠形态学分析，灵敏度不高。一些肿瘤血清学标志也常用于辅助诊断、判断疗效和预后，但所有血清学肿瘤标志物都不是某一种癌症所特有，而是多种都可出现，再加之来自于血清，无组织定位特异性，且有些指标出现相对较晚，这些因素都不利于临床上对肿瘤的特异性诊断和早期发现，给肿瘤的早期诊断带来困难。在肿瘤的演变和恶性转化过程中，基因改变可能潜在出现，比任何形态学变化都早。因此，最早的基因改变可以作为筛选和早期诊断癌症的指标。为鉴定基因突变，建立了以快速PCR技术为基础的新的分子生物学技术，它具有检测10^4个正常细胞中1个突变细胞的潜在能力，故可以用于病人各种体液中突变基因的检测，不需取病人活组织，因而可进行非创伤性检查，而且灵敏度高，假阳性率低，对癌症的早期诊断具有极为重要的意义。

四、PCR技术在遗传病早期诊断中的应用

遗传病是由于人体生殖细胞或受精卵的遗传物质发生改变而引起的疾病，从亲代传至后代，通常包括三大类：①单基因遗传病；②染色体遗传病；③多基因遗传病。遗传性疾病常具有先天性、终生性和家族性的特点。

绝大多数的遗传病目前还没有有效的治疗方法，因此预防就显得尤为重要。若能在妊娠早期诊断出胎儿是否患病，对患病胎儿进行人工流产，即可防止患儿出生。对高风险妊娠进行产前诊断，辅以人工流产，已证明是一种有效预防严重危害健康的遗传病的措施。

与传统的疾病诊断方法一样，遗传病的诊断方法也经历了三个阶段：一是临床学诊断，二是血清学诊断，三是生物化学诊断。以上诊断方法都是以疾病的表型改变为依据的。

现已知道，表型的改变在很多情况下不是特异的，出现的时间往往比较晚，因此有时不能明确诊断，甚至延误病情。随着分子生物学技术的迅速发展和在遗传病研究中的广泛应用，又兴起了第四代诊断技术——基因诊断技术。基因诊断技术的发展是现代医学发展的一个重要标志，它从基因水平开辟了诊断学新领域。聚合酶链反应（PCR）技术是基因诊断主要技术之一，以PCR为基础的各种分子生物学新技术在遗传病的基因诊断等方面得到广泛的应用。

由于PCR技术具有高特异性、高灵敏度、操作快速、简便、对样本要求低等特点，加之它能与多种分子生物学技术配合使用，PCR技术日益成为遗传病诊断、发病机理的研究等方面最有效、最可靠的方法之一。PCR技术的出现为快速、准确地检测人类遗传病开辟了新途径。目前为止文献的报道PCR诊断遗传病种类见表3-3。

表 3-3　目前可用 PCR 检测的人类遗传病种类

急性间歇性卟啉病	无纤维蛋白原血症	遗传性果糖不耐受症	鸟氨酸氨甲酰转移酶缺乏症
腺苷脱氨酶缺乏症	皮肤弹力过度Ⅳ综合征	遗传性持续性胎儿血红蛋白症	苯丙酮尿症
α-1 抗胰蛋白酶缺乏症	皮肤弹力过度Ⅶ综合征	次黄嘌呤磷酸核糖转移酶缺乏症	视网膜母细胞瘤
淀粉样变多发性神经病	椭圆形细胞增多症	低磷酸酶血症	镰刀型贫血
抗凝集素Ⅲ缺乏症	线粒体性脑性肌病	抗胰岛素 A 型病	Sandhoff 病
载脂蛋白 B 缺乏症	家族性高胆固醇血症	Laron 侏儒症	脊柱骨骺发育不全
载脂蛋白 C 缺乏症	葡萄糖-6-磷酸酶缺乏症	Lesch-Nyhan 综合征	Tay-Sachs 病
载脂蛋白 E 缺乏症	Gaucher's 病	白细胞聚合缺乏症	地中海贫血
Becker 肌营养不良症	Gerstmann-Straussler 综合征	枫糖尿症	von^{r} Willebrand 病
囊性纤维病变	视网膜环状萎缩症	非胰岛素依赖性糖尿病	乙型血友病
Duchenne 肌营养不良症	甲型血友病	成骨不全	肝素辅因子 2 缺乏症

五、PCR 在骨髓移植 HLA-D 位点配型中的应用

人类白细胞抗原（HLA）为白细胞的表面蛋白，它是存在于所有脊椎动物中的主要组织相容复合物（MHC）的适用于人类的命名。HLA 基因定位于 6 号染色体短臂 2 区 1 带（$6p^{21}$），编码具有传递抗原功能的细胞表面糖蛋白。HLA 系统在骨髓移植免疫反应中起主导作用，只有供、受者之间 HLA 相合才能进行骨髓移植，因此，HLA 配型的精确性在很大程度上影响着临床的治疗效果。

HLA 经典的配型方法是通过血清学或混合淋巴细胞培养（MLC）方法分析 HLA 基因的表型。近年来，分子生物学配型法丰富了免疫学分型，它不仅不受机体免疫状态或标本量少（如化疗后）等因素的影响，而且可以直接分析基因型。聚合酶链反应（PCR）技术的问世，使得人们能够按照自己的需要获得大量的特异性扩增的 DNA 序列，序列特异的寡核苷酸探针杂交（sequence specific oligonucleotide probe hybridization，SSOPH）可直接检测样本的核苷酸序列，是检测多态性的最好方法。PCR 技术与 SSOPH 结合甚至可检测出 HLA-D 区单个碱基改变的微多态性，其步骤包括：①PCR 扩增 HLA-D 区具有多态性的特异性序列；②固定放大产物于尼龙膜或硝酸纤维素膜上；③标记的寡核苷酸探针与滤膜杂交；④用同位素自显影、酶显色或发光法检测结合的探针，确定基因类型。目前 PCR；SSOPH 方法已用于骨髓移植配型、法医学中个体鉴定以及疾病相关性的研究。

为了尽快选择供者，不延误骨髓移植，常要求尽可能快地进行配型。最近建立了一种简单、快速的 HLA-DRβ 位点配型方法，即 PCR 指纹图，该技术是将供者、受者的 DNA 经 PCR 扩增 HLA-DRβ 区基因的高度多态区域，在分别扩增后将供、受者产物混合，混合后进行 2 个 PCR 循环，扩增产物经 12%的聚丙烯酰胺凝胶电泳，照相后分析不同个体的指纹图型。当 HLA-DRβ 基因相合时，其基因产物在聚丙烯酰胺凝胶中可产生一致的带型，这种多条带型的产生是由于各种 HLA 单倍体中的 HLA-DRβ 基因的不同亚型以及假基因上的多态性所致。当 PCR 最后两个匹配循环退火时，这些基因的单链 DNA 复性，DRβ 基因相同的个体形成同型双链，而不同的 DRβ 基因之间的不同序列形成异型双链，因此，混合匹配可进一步证实 HLA-DRβ 是否相合。

六、PCR 技术在法医学鉴定中的应用

随着分子生物学及分子遗传学的深入研究，重组 DNA 技术、DNA 扩增（PCR）技术的建立及应用，法医学 DNA 分析技术应运而生，法医物证检验已进入 DNA 水平。PCR 技

术以它特有的高灵敏度、高特异性和快速的反应时间为 DNA 多态性分析提供了有力的工具，可以对犯罪现场或个体的检查中通常收集到的各种形式的生物学证据，如牙齿、骨骸、血液、精斑、单根毛发，甚至单个精子的 DNA 加以检测分析。PCR 技术用于生物物证的分析，与 RFLP 技术相比，具有操作简单、省时、成本低廉的优点，还可避免同位素分析所带来的一系列问题，如放射性污染、操作时的人身防护以及材料来源困难等。近年来，在 PCR 技术基础上，又发展了许多基因位点分型系统，用于法医物证的 DNA 分析，其中发展最为完善并得到公认的是 HLA-DQa 位点的分型系统，此系统应用 PCR 技术扩增 HLA-DQa 基因特异片段，结合等位基因特异性寡核苷酸探针杂交技术检测 HLA-DQa 位点的 4 个等位基因，有特异性好、灵敏度高、结果稳定可靠等优点，成为国际上法医科学中应用 PCR 技术进行个人识别和亲子鉴定最有效的方法。

人类高度可变 DNA 标记的发现和应用，给传统的法医学方法带来了一次冲击。多数 DNA 分析系统采用串联重复的小卫星 DNA，或采用由于串联的重复单位数目变化而引起等位基因变化的 VNTR 系统。用 PCR 技术扩增高可变位点，包括由非常短小串联重复结构形成的微小卫星（microsatellite），可以极大地提高 DNA 分析的灵敏度，增加检测降解 DNA 的能力。

扩增片段长度多态性是指选择与 VNTR 位点两端互补的序列来合成引物，然后用 PCR 扩增长度不同的片段，并用高分辨率的简单小型聚丙烯酰胺凝胶电泳来区分，最后用硝酸银染色得到结果。该法分辨率、灵敏度明显高于使用琼脂糖凝胶电泳、EB 染色的分析方法，并且操作时间短，一般 2 h 内即可完成，使 VNTR 在琼脂糖电泳中连续分布的等位基因片段变成单独分离的片段，使原来无法用 Hardy-Weinberg 公式进行计算的基因频率估计成为可能，使其在法医学上的应用日益广泛。

总之 PCR 技术的应用，给法医科学带来了一场深刻的变革，展示出广阔的应用前景，并将成为有效的物证检验手段。从搜查犯罪现场中收集的或从罪犯身上取得的所有的生物学证据均适用于 PCR 分析。因而，精液、血液污染物、单根头发、口腔上皮细胞都可进行 DNA 分型。从理论上讲，只要有一个包含靶序列的 DNA 分子，甚至微量降解的 DNA，也可以进行分析。

七、PCR 在进化分析中的应用

rRNA 基因有“化石”之称，PCR 扩增并分析其序列，可分析比较界或门分类水平上的生物。通常，细胞核小亚基 rDNA 的序列进化相对较慢，主要用于种系关系较远的生物比较研究；而线粒体 rRNA 基因进化非常快，可在科目水平上进行不同生物的比较；转录的间隔区序列和核 rRNA 的基因间隔区重复单位序列进化最快，可用于同属异种或不同群之间的分析比较。

（严世荣　刘志国）

第四章　分子杂交与印迹技术

核酸分子杂交技术是基于核酸分子的变性和复性原理基础上产生和发展起来的。杂交的一方是待测的 DNA 或 RNA，另一方是已知的 DNA 或 RNA 的片段。核酸分子杂交技术是用已知碱基序列的单链核苷酸片段作为探针去探测样品中是否存在与其互补的同源序列的一种检测方法。这种检测方法具有高度的特异性和灵敏性，是定性或定量检测特异 DNA 或 RNA 序列片段的方法之一。核酸杂交及在此基础上发展的各种印迹技术被广泛地应用于分子生物学领域中克隆基因的筛选、DNA 序列测定、基因组中特定基因序列的检测、基因突变分析以及疾病的诊断等方面。本章将介绍核酸杂交的基本理论、核酸探针的标记以及利用分子杂交的各种印迹技术。

第一节　核酸杂交的基本理论

一、DNA 变性、复性及杂交

1. DNA 变性

DNA 变性（DNA denaturation）是指在物理或化学因素的作用下，DNA 分子互补碱基对之间的氢键断裂，使双链 DNA 分子变成两条单链 DNA 的过程。变性过程中，只是链间的氢键断裂，而核酸分子中的所有共价键（如磷酸二酯键、糖苷键等）则不受影响。

引起核酸变性的因素有很多，常见的有加热、酸、碱、乙醇、丙酮、尿素、甲酰胺、甲醛等。如加热变性是将温度升高到 90℃以上，此时 DNA 氢键完全断裂成为单链；酸碱变性是将核酸溶液的 pH 调至 3 以下或 10 以上时，可使核酸的双链完全打开成为单链核酸分子；化学试剂变性是将一定浓度的化学试剂（如尿素、甲酰胺、甲醛等）加入到核酸溶液中，使两条链之间的氢键断裂形成单链的过程。加热变性是核酸分子杂交中最常用的一种方法。

双链 DNA 变性一半所需要的温度称为 DNA 的溶解温度（melting temperature，T_m），一种 DNA 分子的 T_m 值的大小与所含碱基中的 GC 比例有关。由于 GC 之间是三个氢键，AT 之间是两个氢键，所以当 DNA 中 GC% 愈高时，T_m 值也就愈高。T_m 值还受溶液中离子强度的影响，离子强度低，T_m 值就低。如单价离子浓度增加 10 倍，则 T_m 值约增加 16.6℃，这是由于溶液中离子与 DNA 分子中磷酸基团形成了离子键；因此离子强度高，DNA 比较稳定，需要更多的能量才能使其变性。甲酰胺可使碱基对间的氢键不稳定，因此 DNA 在甲酰胺溶液中的 T_m 可随甲酰胺浓度增高而降低，这可避免 DNA 在高温变性时引起断裂，因而在分子生物学实验中常常被应用。

2. DNA 复性

DNA 的变性是可逆的。当促使变性的因素解除后，两条 DNA 链又可通过碱基互补配

对结合形成双链 DNA 分子，这一过程称为 DNA 复性（renaturation）。加热变性的 DNA 分子在温度缓慢降低时可恢复到原来正常 DNA 的结构，这一复性过程又称为“退火”（annealing）。如果将加热变性的 DNA 分子快速冷却至低温，则大部分 DNA 不能复性。

复性过程并不是两条单链重新缠绕的简单过程，而是从单链分子的无规则碰撞运动开始的，这种碰撞过程是随机的。首先，两条单链核酸随机碰撞形成局部的双链，所形成的局部双链是暂时的，如果此局部双链周围的碱基不能配对，则所形成的双链会重新解离，继续随机碰撞。当随机碰撞找到了正确的互补顺序，那么，形成的局部双链就成为中心顺序，其两侧的顺序会迅速互补配对，形成完整的双链分子。

如果促使 DNA 变性的因素过分强烈，例如加热到 100℃以上或 pH 低于 3.0，将导致 DNA 分子中共价键的断裂，DNA 分子便难以复性。

3. 核酸分子杂交的概念

将任何具有互补核苷酸顺序的两条单股核酸分子放入同一反应体系，两条互补链可通过复性重新缔合形成双链，这一过程称为核酸分子杂交（hybridization）。核酸分子杂交可以发生在 DNA/DNA、DNA/RNA、RNA/RNA 或人工合成的寡核苷酸片段与 DNA、RNA 等片段之间。核酸分子杂交是分子生物学常用的重要技术之一。利用核酸分子杂交可以检出特定基因的顺序、基因组织的特点、基因的结构和定位、基因的表达等。

二、核酸探针

探针（probe）是指带有标记物的、能与被检测的核酸片段互补的一段已知核酸片段。探针可以是用基因克隆技术分离获得的特异 DNA 序列，或是特异 DNA 序列在体外转录出的 RNA 序列或 cDNA 序列，也可是人工合成的寡核苷酸片段。理想的探针必须具备以下几点：①探针必须是一段已知的核苷酸序列；②必须加以标记；③便于杂交后的检测。根据核酸探针的不同来源可将探针分为基因组 DNA 探针、cDNA 探针、寡核苷酸探针和 RNA 探针。

1. 基因组 DNA 探针

从基因组中获得的几乎所有的基因片段都可以被克隆到质粒或噬菌体载体中，通过大量扩增、纯化，酶切 DNA 片段，分离纯化，标记后即可作为探针使用。基因组 DNA 探针可以是基因组全序列（如较小的 DNA 病毒探针），可以是某一个特定的基因，也可以是基因上的一个片段。

2. cDNA 探针

首先从相应的真核细胞中分离获得总 mRNA，经逆转录获得 cDNA 后，将其克隆于适当的克隆载体，通过扩增重组质粒而使 cDNA 得到大量扩增。提取质粒后，用限制性内切酶消化重组质粒，将 cDNA 片段切割下来，分离纯化，标记后即可作为探针使用。与基因组 DNA 探针相比，cDNA 探针不存在内含子或其他高度重复序列，是目前应用最为广泛的一种探针。

3. 寡核苷酸探针

采用 DNA 合成仪合成一定长度的已知寡核苷酸片段，经标记后也可作为探针使用。其优点是可根据需要随心所欲地合成相应的序列，避免了天然核酸探针中存在的高度重复序列所带来的不利影响。对于那些未知的核酸顺序，可根据蛋白质的氨基酸顺序推导出核酸顺序，合成相应的寡核苷酸片段作为探针进行探测。此种情况下需要考虑到密码子具有简并性；采用推导的序列制备探针，也存在其特异性可能不高等问题。寡核苷酸探针多用于克隆

筛选和点突变分析。

4. RNA 探针

RNA 探针可通过 SP6 和 T7 体外转录体系制备。将目的基因克隆到带有 SP6 或 T7 启动子的质粒中，插入到启动子下游，用适当的限制性内切酶在插入序列的下游切断质粒，使之线性化，加入 RNA 聚合酶、ATP、GTP、CTP 和 [α-^{32}P]-UTP，以目的基因的 DNA 为模板，转录合成出高放射活性的 RNA 探针。

RNA 探针的优点在于：RNA 探针为单链分子，杂交时不存在互补双链的竞争性结合，杂交效率高，杂交体较稳定，非特异性杂交较少以及本底低等。

三、核酸杂交方法

核酸分子杂交可采用两种方法：固相杂交法和液相杂交法。固相杂交是把待检测的核酸样本首先结合到某种固相载体（如硝酸纤维滤膜、尼龙膜等）上，然后将其置于反应液中，与核酸探针进行杂交反应。根据支持物的不同，固相杂交又分为两种类型：原位杂交和滤膜杂交。前者是指不改变核酸原来的位置，探针直接与细胞或组织切片中的核酸进行杂交；后者是指探针与固定于滤膜上的核酸分子进行杂交。液相杂交是指待测的核酸和标记的探针同时溶于杂交液中进行反应，然后分离杂交双链和未参加反应的标记探针，再进行检测。核酸杂交是一个较复杂的反应过程，受很多因素的影响。下面就其影响因素和可以采用的两种杂交方法作一简单介绍。

（一）影响杂交的因素

1. 探针的浓度和长度

探针浓度的选择是以达到与靶核酸的最大饱和度为目的，一般选择 0.5～5.0μg/ml。探针浓度的高低直接影响单链核酸间的碰撞概率和杂交反应速度。浓度大，碰撞概率高，杂交反应速度快；浓度小，碰撞概率低，杂交反应速度慢。另外，反应速度的快慢与探针片段的长度也有关，探针片段越长，扩散速度越慢，核酸复性的速度也越慢。探针的长度一般应控制在 50～300bp。

2. 杂交的温度和时间

核酸杂交反应受杂交温度的影响，在 0℃时杂交进行非常慢，随着温度的升高，杂交率也明显增加，当温度比 T_m 值低 20～25℃时达到最大杂交率（DNA-DNA 杂交）。但在更高温度情况下，双链分子逐渐趋向解链，故当温度达到比 T_m 低 5℃时杂交率非常低。所以，杂交温度一般在比 T_m 低 20～25℃。

杂交反应时间一般为 16h，要注意不要超过 24h。反应时间长，形成的杂交体会自动解链，杂交信号反而减弱。

3. 杂交液中的甲酰胺

甲酰胺能降低核酸杂交的 T_m，含 30%～50%甲酰胺的杂交溶液其杂交温度能降低到 30～42℃。使用甲酰胺具有以下优点：①在低温下探针更稳定；②能更好地保留非共价结合的核酸。在实际工作中，对于待测核酸顺序与探针顺序同源性高的杂交，用水溶液时取 68℃；用 50%甲酰胺溶液时常在 40℃进行杂交。如果待测核酸顺序与探针同源性不高时，以 50%甲酰胺溶液在 35～42℃杂交为好。

4. 离子强度

在低离子强度下，核酸杂交非常缓慢，随着离子强度的增加，杂交反应率增加。例如，

在低盐浓度（<0.1mol/L Na^+）时杂交率较低，当盐浓度每增加 2 倍时杂交率增加 5～10 倍，盐浓度>0.1mol/L Na^+ 时对杂交率的影响降低。高浓度的盐使碱基错配的杂交体更稳定，故测定顺序不完全同源的交叉杂交时，必须维持杂交反应液中较高的盐浓度和洗膜溶液的盐浓度。杂交后一般用较高浓度的盐溶液洗膜，如 2×SSC 和 6×SSC（1×SSC：150mmol/L NaCl，15mmol/L 柠檬酸钠）。

5. 非特异性杂交反应

为减少非特异性杂交反应，在杂交前将非特异性杂交位点进行封闭，以减少其对探针的非特异性吸附作用。常用的封闭物有两类：一类是变性的非特异性 DNA，大多采用鲑鱼精子 DNA（salmon sperm DNA）或小牛胸腺 DNA（calf thymus DNA）；另一类是一些高分子化合物，一般多采用 Denhardt 溶液（含聚蔗糖 400、聚乙烯吡咯烷酮和牛血清白蛋白），也可用脱脂奶粉。

6. 杂交促进剂

杂交液中使用葡聚糖可促进核酸链间碱基的配对，同时其微粒表面还可吸附探针分子，从而使核酸单链接触面积增大，特别有利于长探针的杂交反应，在 5%～10%硫酸葡聚糖存在时，杂交反应速率可提高 5～10 倍。使用硫酸葡聚糖的缺点是增加了杂交液的黏度。聚乙二醇（PEG，M_r 6000～8000）和聚丙烯酸也可作为杂交促进剂，这二种物质具有黏度低、价廉等优点，但实际工作中可根据情况选用。

（二）固相杂交法

1. 固相支持膜的选择

进行固相核酸分子杂交时，选择良好的固相支持膜是核酸分子杂交成败的关键因素之一。用于核酸杂交的膜应当具备的条件是：①有较强的结合核酸分子的能力；②核酸固定在膜上并不影响核酸与探针的杂交反应；③能经受杂交、洗膜等反复操作而杂交分子不易脱落；④非特异性分子吸附少而不牢固，易于洗脱；⑤具有良好的柔软性和韧性，便于实验操作。下面介绍几种固相支持膜。

（1）硝酸纤维素滤膜（nitrocellulose filter membrane）　硝酸纤维素滤膜与核酸的结合有赖于高盐浓度（>10×SSC），在高盐浓度下，该膜具有较强的吸附单链 DNA 和 RNA 的能力。吸附的单链 DNA 和 RNA 经真空中 80℃烤干后，依靠疏水性相互作用而结合到硝酸纤维素滤膜上。硝酸纤维素滤膜具有杂交信号本底较低以及非特异性吸附蛋白质的作用弱等优点，因而被广泛应用于 Southern 印迹、Northern 印迹、斑点印迹以及涉及蛋白质作用的非放射性标记探针的杂交体系。硝酸纤维素滤膜的缺点是在低盐浓度时结合 DNA 效果不佳，不适宜于电转印迹法，对于小分子量 DNA（<200bp）的结合能力不强，质地脆弱，不易操作等。

（2）尼龙膜（nylon membrane）　与硝酸纤维素滤膜相比，尼龙膜结合单链及双链 DNA 和 RNA 的能力要强的多，是目前较为理想的一种核酸固相支持膜。经烘烤或紫外线照射后，核酸分子与膜共价结合；特别是用短波紫外线照射后，核酸中的部分嘧啶碱基可与膜上带正电荷的氨基酸相互交联，从而使结合更加牢固。尼龙膜可与小至 10bp 的片段结合，在低盐条件下也能较好地结合核酸。该膜韧性较强，操作方便，可重复使用。一次杂交后，探针分子可经碱变性而被洗脱下来，因而可用于与第二探针进行杂交。尼龙膜的缺点是杂交信号本底较高，但可采用加大预杂交液中非特异性封闭试剂的方法克服。

（3）化学活化膜（chemical activated paper）　用一定的化学物质处理滤膜后，使膜上产

生某种活性基团的滤膜即为化学活化性膜。膜上的活性基团（如重氮盐），可以与 DNA 或 RNA 分子共价结合。与硝酸纤维素滤膜和尼龙膜相比，该膜的特点是对不同大小的核酸片段都具有同等的结合能力。但是化学活化膜由于活化过程较复杂，结合能力较硝酸纤维素滤膜要低，因而较少被使用。

2. 常用的几种固相核酸分子杂交方法

（1）Southern 印迹杂交（Southern blot） 是指将电泳分离、原位变性的单链 DNA 片段转移到固相支持物上，然后与核酸探针杂交的一种方法。该方法的基本步骤包括：待测核酸的制备→限制性内切酶的消化→琼脂糖凝胶电泳分离→原位 DNA 变性→DNA 转移至硝酸纤维素滤膜→预杂交→杂交→洗脱→放射自显影→分析杂交结果。由于转移到滤膜上的各个 DNA 片段的相对位置与在凝胶中的相对位置完全一样，因而称为印迹（blot）。

（2）Northern 印迹杂交（Northern blot） 是指将待测 RNA 样品经变性电泳凝胶分离后转移到固相支持物上（如尼龙膜等），然后与标记的核酸探针进行杂交的一种方法。该方法主要用于组织细胞中总 RNA 或 mRNA 的定性和定量分析；可根据 mRNA 分子在电泳中的迁移位置与分子质量标准参照物比较，判断待测 RNA 样品分子质量的大小；对杂交信号的强弱比较，反映该基因转录 mRNA 水平的高低。Northern 印迹杂交是广泛用于基因表达调控、基因结构与功能、遗传变异及病理研究的一种方法。

（3）斑点印迹及狭缝印迹 将 RNA 或 DNA 变性后直接加样于多孔过滤加样器并以固定的带型聚集于硝酸纤维素膜或尼龙膜上，膜上的印迹为圆形，称斑点印迹（dot blot）；印迹为线状，称狭缝印迹（slot blot）。然后，用标记的核酸探针与之杂交。也可将各种不同的探针先固定于滤膜上，再与样品进行杂交，这种方法称为反向杂交。斑点及狭缝印迹杂交都可采用放射自显影或显色反应检测杂交信号，分析杂交结果，可用于基因组中特定基因及其表达的定性及定量研究。这两种杂交方法的区别只在于点样的形状不同，而操作方法基本相同。

（4）菌落原位杂交（colony hybridization in situ） 是将生长在平板上的菌落，影印转移到膜上，经裂菌、DNA 变性、中和等处理后，进行杂交的过程。该法可用于基因工程研究中 DNA 重组体的筛选和临床标本的检测。

（5）核酸原位杂交（nucleic acid hybridization in situ） 是通过适当方法处理细胞或组织后，使核酸保持在细胞或组织切片中，用标记的核酸探针与其进行杂交的方法。原位杂交不需要从组织或细胞中提取核酸，对组织中含量极低的靶序列有很高的灵敏度，并可完整地保持组织与细胞的形态，能准确地反映出组织细胞的相互关系及功能状态。根据检测物不同，原位杂交可分为细胞内原位杂交和组织切片原位杂交。

（三）液相杂交法

液相杂交时要根据探针的种类和待测核酸种类不同，设计反应条件，采用不同的杂交双链的分离方法。由于实验的目的不同，反应条件和操作方法也有一定的差异，液相杂交不能简单地确定一个固定的条件和程序。目前常用的液相杂交有核酸酶 S1 保护分析法、RNA 酶保护分析法。

1. 核酸酶 S1 保护分析法

核酸酶 S1 保护分析法（nuclease S1 protection assay）是一种比 Northern 印迹杂交敏感性更高的检测 RNA 的方法。将基因组 DNA 标记成为探针，与待测 RNA 在适当的液相条件下进行杂交，形成 DNA/RNA 杂交双链。在杂交后的反应体系中加入核酸酶 S1 进行处

理，未与 RNA 结合的单链 DNA 被降解，DNA 中的内含子部分不能与 RNA 结合而形成游离环状，也被核酸酶 S1 降解。利用核酸酶 S1 能专一性地降解单链 DNA 和 RNA，而 DNA/RNA 杂交双链则受到保护的特点，通过聚丙烯酰胺或琼脂糖凝胶电泳分析消化产物，即可确定不被降解的 DNA 片段大小，此法可用于基因转录起始位点分析及内含子剪切位点分析（图 4-1）。

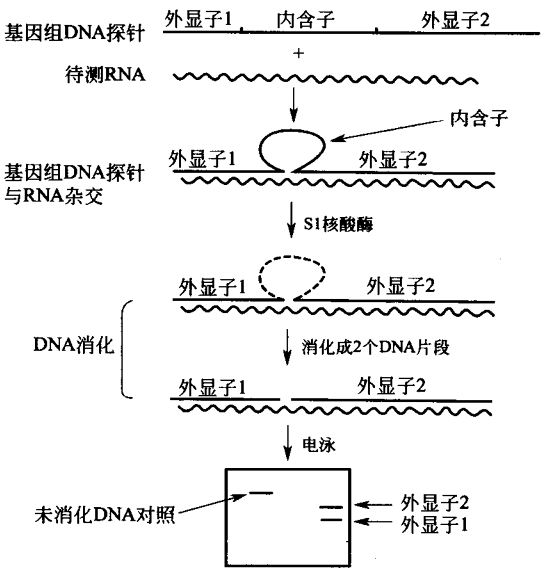


图 4-1　RNA 的 S1 核酸酶保护分析

2. RNA 酶保护分析法

RNA 酶保护分析法（RNase protection assay，RPA）是利用 RNase A 和 RNase T1 能专一性降解单链 RNA，而双链 RNA 则受到保护的特点，用体外转录合成的放射性标记的 RNA 探针与待测 mRNA 进行液相杂交，使互补序列 RNA 探针和待测 RNA 形成杂交体进行检测的方法。

（1）基本特点　RPA 与传统的方法比较具有特异性强、灵敏度高、操作简便、周期短、检测量大、准确性高等优点。相对于 RT-PCR 而言，RPA 没有扩增过程，因而减少了误差，提高了重复性；相对于 Northern blot 而言，RPA 没有转膜和杂交后洗膜的过程，采用的是液相杂交，基本上没有样品损失，提高了检测灵敏度，RPA 比 Northern blot 灵敏 15～150 倍，可以检测出 0.1pg 的 mRNA，适合检测各种表达水平的基因；相对于原位杂交而言，RPA 没有组织制片过程，因而操作简单、周期短。由于在 RNA 酶消化过程中，与探针错配的 RNA、单链探针和未杂交 RNA 都将被水解，而仅保留特异性的 RNA-RNA 杂交分子，因而 RPA 具有较高的特异性。从该方法的设计思路上看，探针是与目标 mRNA 中部分片段杂交，杂交片段长度基本上等同于探针长度，而不像 Northern blot 是探针与 mRNA 全长片段杂交，因此 RPA 的方法可以用于部分降解的 RNA 样品检测，这是 RPA 的又一优点。根据实践，从 RNA 提取、探针制备到放射自显影，最短可以做到 4d 出结果。

RPA 的另一大优点是可以将若干个探针放入同一个体系内杂交，不存在竞争性干扰，使一次操作可以得到多个数据；由于 RPA 的高通量，同时检测多个基因，可评价它们在同一情况下的差异表达；一次 RPA 就能完成十多次 Northern blot 的工作，大大加快了研究速度和数量。由于用 RNase 水解 RNA 杂交反应中单链 RNA 产生的假带较用 S1 水解 RNA/DNA 反应中单链核酸产生的假带少，所以该方法逐渐成为检测 RNA 的一种标准化方法。根据待测 RNA 的特性，设计不同的 RNA 探针，可以达到不同的目的。下面介绍 RPA 的基本操作流程。

（2）RPA 流程　RPA 操作程序见图 4-2。

① 待测细胞总 RNA 的提取　异硫氰酸胍法、Trizol 试剂法、oligo(dT) 纤维素层析法提取总 RNA，根据总 RNA 中 mRNA 的含量，确定 RPA 中总 RNA 的用量。一般为 30～50μg。

② RNA 探针的体外转录标记　体外转录标记 RNA 探针具有几个优点：一是用来合成 RNA 的模板可以在反应体系中多次被转录，因此产生的放射性 RNA 量远远超过模板的量，

使标记效率高；二是 RNA 体外转录所用的 DNA 模板可以用 DNase Ⅰ消化除去，不需电泳纯化 RNA 探针；三是 RNA 探针产生的杂交信号一般比 DNA 探针强。

a. 体外转录模板 DNA 的制备　将模板 DNA 片段克隆入含有 T7、SP6 或者 T3 启动子的克隆载体的多克隆位点中，使 DNA 片段位于启动子的下游，并确定双链 DNA 模板中哪条链用作 RNA 转录的模板，然后用适当的限制性内切酶在 DNA 片段的下游线性化重组质粒 DNA。接着用酚-氯仿抽提上述酶切产物，乙醇沉淀，用 TE 缓冲液（pH7.5）溶解模板 DNA。

b. 体外转录标记 RNA 探针　将转录体系加入离心管中，在 T7（SP6 或 T3）RNA 聚合酶作用下，带有同位素标记的 ^{32}P-UTP 掺入到转录的 RNA 探针中。

c. RNA 探针的回收和纯化　在上述混合反应物中加入无 RNase 污染的 DNase Ⅰ，混合后 37℃温育 15min。然后酚-氯仿抽提，最后将 RNA 溶于无 RNase 污染的 DEPC 水中，此 RNA 即可用作探针。若不立即使用，应在无水乙醇中－70℃保存。

③ 待测 RNA 与 RNA 探针的液相杂交

a. 待测 RNA 样品的准备　将乙醇沉淀的细胞总 RNA 溶于 DEPC 水中，然后空气蒸发到 RNA 刚好干为止。用杂交缓冲液溶解 RNA 样品，放置冰上待用。

b. 液相杂交反应　将同位素标记的 RNA 探针与细胞总 RNA 混合，无水乙醇沉淀，然后将 RNA 沉淀溶解在杂交缓冲液中，置 90℃水浴 10min，迅速移入 40℃水浴中杂交 12h。最后 RNase 消化。

④ 酚-氯仿抽提待测 RNA 和 RNA 探针的杂交体　按常规方法抽提。

⑤ 聚丙烯酰胺凝胶电泳及放射自显影　将样品加入 5%聚丙烯酰胺凝胶的加样孔中，同时设标记的 RNA 探针为阳性对照，RNase 消化后的酵母 tRNA 为阴性对照，并设标准参照物。

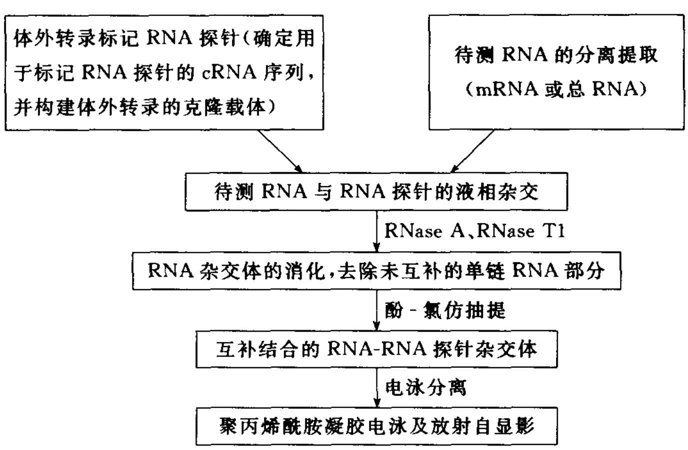


图 4-2　RPA 操作程序示意图

（3）注意事项

① RNA 探针的标记采用的是体外转录法，RNase 对 RNA 的降解作用是影响 RNA 质量从而影响 RNase 保护实验结果的重要因素，因此需尽可能排除 RNase 的污染。

② RNA 样品的制备对于实验的成败非常关键，如果制备的 RNA 样品中 mRNA 不完整，可以造成结果的多样性或假阳性，所以应采用快速法提取细胞 RNA，尽可能缩短对 RNA 的操作时间。

③ RNA-RNA 杂交体的电泳采用变性胶，所以看到的结果是探针所显示出的片段。

④ 杂交反应中，需设置针对所有探针的阳性及阴性对照，排除交叉杂交的可能性。

（四）杂交信号的检测

核酸探针标记的方法有两大类：放射性同位素标记和非放射性标记物标记。两种标记的核酸探针与待测的单链核酸分子的杂交过程基本相似，但杂交信号检测体系是完全不同的。

1. 放射自显影

放射性同位素探针杂交信号一般是采用放射自显影的方法来检测。经杂交、洗脱后的膜放入带有增感屏的曝光暗盒内，在暗室内将一张 X 光底片放入曝光暗盒，使之接触滤膜。再将暗盒放入－70℃冰箱曝光一定时间。在此间，同位素在不断衰变的过程中释放出 β 粒子，粒子撞击 X 光片感光层，形成潜在影像。从冰箱取出暗盒，使其温度上升至室温，经过显影即可见到成像。放射自显影时用增感屏可提高检测灵敏度，高能的 β 粒子除直接撞击 X 光感光层使之感光外，粒子撞击含无机磷的增感屏发出荧光使 X 光片进一步感光，放射自显影的效果可增加 10 倍。低温下可减缓溴化银结晶在光子激活后回复到稳态的速度，也能增加感光效果。X 光片曝光时间长短可根据所用放射性核素的种类和含量来决定，也可参考预实验结果。影像的分辨率高低则取决于射线类型和放射性同位素能量大小。

2. 非放射性探针的检测

除酶直接标记的探针外，大多数非放射性标记物是半抗原，不能被直接检测，需经偶联及显色二步反应进行检测。

地高辛（Dig）是最常用和商品化的非放射性标记物。可采用随机引物法（切口平移法和末端转移酶末端标记法等），在 Klenow 酶的作用下，将地高辛标记的 dNTP 作为底物掺入到 DNA 中，合成标记 DNA 探针。然后，与固定在滤膜上的同源 DNA 杂交，杂交信号可通过酶促显色反应显示。在反应体系中，加入抗地高辛抗体-碱性磷酸酶（AP）复合物。此时，抗地高辛抗体与探针分子中的地高辛形成半抗原-抗体结合物。结合在抗体上的碱性磷酸酶（alkaline phosphatase）可使其作用底物 5-溴-4-氯-3-吲哚磷酸（5-bromo-4-chloro-3-indoly phosphate，BCIP）脱磷并聚合，在此过程中释放出的 H^+ 使硝基蓝四氮唑（nitroblue tetrazolium，NBT）还原，形成紫蓝色化合物，根据显色反应判断实验结果。

很多荧光染料和碱性磷酸酶、辣根过氧化物酶（产生有颜色的沉淀物）可直接与抗地高辛抗体和链霉亲和素偶连。现已证实链霉亲和素-碱性磷酸酶-BCIP/NBT 联合应用检测生物素标记探针杂交信号敏感性最高。化学发光检测方法和间接免疫荧光技术也为许多分子生物学的应用提供了另一灵敏的备选方法。

第二节　常用印迹杂交方法

一、Southern 印迹杂交法

Southern blot 杂交是 1975 年由爱丁堡大学的 Southern 建立的，由此得名“Southern blot”。它是将电泳分离、原位变性的单链 DNA 片段转移到固相支持物上，与核酸探针杂交

后检测的一种方法。Southern blot 杂交技术可用于基因组中特定基因的定性和定量分析、基因突变和缺失的分析及限制性片段长度多态性分析等。Southern blot 在疾病的诊断和治疗等方面也得到广泛地应用。Southern blot 从核酸样品制备、酶切消化、电泳分离到 DNA 原位转移至硝酸纤维素滤膜、杂交检测等经过十多个步骤，如图 4-3 所示。其主要步骤分述如下。

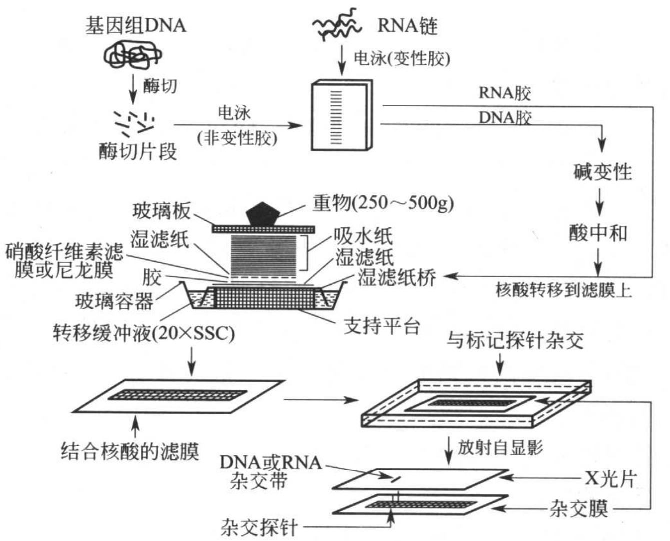


图 4-3　核酸转模、杂交基本过程示意图

1. DNA 样品的制备

DNA 样品可以从动物组织细胞或培养细胞中制备，也可以从克隆载体中获得。用一种限制性内切酶或分别用不同的限制性内切酶消化 DNA 样品，得到一定分子质量大小的 DNA 片段混合物。将不同大小的 DNA 样品加入在 0.8%～1.0%非变性琼脂糖凝胶中进行电泳，将大小不同的 DNA 片段分离开来。推荐的样品 DNA 上样量列于表 4-1 中。

表 4-1　推荐使用的上样量

样品 DNA	上样量(单个泳道的 DNA 样品)	使用探针的类型
基因组 DNA	1.0μg、2.5μg、5.0μg(3 个不同样品)	Dig 标记的 DNA 或 RNA
质粒 DNA	<1ng	Dig 标记的 DNA 或 RNA
总 RNA	≤1μg	Dig 标记的 RNA
	≤5μg	Dig 标记的 DNA
mRNA	≤100ng	Dig 标记的 RNA
	≤500ng	Dig 标记的 DNA

DNA 样品在制备和电泳过程中始终保持双链结构。电泳后，将凝胶置于碱性溶液中浸泡，使双链 DNA 变性成为单链 DNA，再用酸中和，使溶液的 pH 恢复至中性。然后，将凝胶中的 DNA 片段转移到硝酸纤维素滤膜上。

2. DNA 样品的转膜

DNA 样品转膜常见的有 3 种方法：虹吸印迹法，电转移法，真空转移法。3 种方法在许多方面都存在差异，其中较为常用的是虹吸印迹法。有关 3 种方法的差别列于表 4-2。

表 4-2　不同印迹方法的比较

方　法	印迹前凝胶的处理	盐离子的量	操作技术	优　点
虹吸印迹法 (capillary transfer)	Southern：脱嘌呤（如果靶基因>5kb），变性，中和 Northern：在 20×SSC 中浸润 10min，去除甲醛	20×SSC 作为转移缓冲液（如果做 Northern 则应无 RNase）	目的基因在缓冲液中单向被动转移	灵敏度高，非常可信，且重复性好，转移完全
电转移法 (eletroblotting)	在转移缓冲液中平衡	按照厂家使用说明	使用电流增加转移速度（仅推荐用于聚丙烯酰胺凝胶）	快速转移
真空转移法 (vacublotting)	与虹吸印迹法相同	对于 DNA 或 RNA 使用 20×SSC	使用真空加快转移速度，通常使用碱转移法（0.4mol/L NaOH 作为转移缓冲液，仅用于 DNA）（注意：有报道显示，使用 Dig 系统时，使用碱转移法会降低敏感性）	快速转移（大约 1h）（注意：不如虹吸印迹法更敏感）

3. 转膜步骤

① 如果靶序列>5kb，则需进行脱嘌呤处理：把凝胶浸在 0.25mol/L HCl 中，室温轻轻晃动，直到溴酚蓝从蓝变黄（注意：处理人类基因组 DNA 不要超过 10min，处理植物基因组 DNA 不要超过 20min）。然后，将凝胶浸在灭菌双蒸水中。

如果靶序列<5 kb，则直接把凝胶浸在变性液（0.5mol/L NaOH，1.5mol/L NaCl）中轻轻晃动，室温浸泡 15min，换液重复 1 次。然后将把凝胶浸在灭菌双蒸水中。

② 把凝胶浸在中和液中 [0.5mol/L Tris-HCl（pH7.5），1.5mol/L NaCl]，室温浸泡 15min，换液重复 1 次。

③ 在 20×SSC 中平衡凝胶至少 10min。

④ 按下列操作步骤进行转膜（避免气泡的形成）：a. 用 20×SSC 浸湿一张 Whatman 3MM 滤纸，放在支撑台上，形成一个“桥”；b. 把凝胶放在浸湿的 Whatman 3MM 滤纸上，注意两者之间不要有气泡；c. 按照凝胶的大小剪一张尼龙膜；d. 把膜放在凝胶上；e. 在尼龙膜上放一张 Whatman 3MM 滤纸，一叠吸水纸，上置一玻璃板，其上压一重约 250～500g 的物品。

⑤ 在 20×SSC 的转移缓冲液中过夜，充分转移。

⑥ 可通过以下方法把 DNA 固定在膜上。

紫外交联：把膜（DNA 面朝上）放在用 2×SSC 浸湿的 Whatman 3MM 滤纸上，暴露在紫外光下 1～3min，然后在无菌双蒸水中浸润膜，最后空气中晾干膜。

干烤：在 2×SSC 中短暂洗膜，随后在 120℃干烤固定 30min 或 80℃干烤固定 2h。

⑦ 如果不立即进行下一步，可把膜保存在两张 Whatman 3MM 滤纸之间，放在密封袋中，4℃保存。

将固定于膜上的 DNA 片段与探针 DNA 进行杂交之前，必须先进行一个预杂交（pre-hybridization）的过程。由于硝酸纤维素滤膜能与 DNA 和一些大分子物质结合，为了防止探针 DNA 与滤膜非特异性地结合，在杂交前，必须将膜上所有能与 DNA 和一些大分子物质（如蛋白质）结合的位点全部封闭，达到使本底降低、背景清晰的目的。

4. 预杂交

杂交膜在预杂交、杂交和检测过程的任何时候都不能干燥，否则检测结果会产生高背景（例外情况：最后一次高严谨性冲洗后，膜可以在空气中干燥，然后保存在 4℃。然而，这样的膜不能再用于探针的剥离和再杂交）。实验步骤如下。

① 按照探针、靶基因和杂交液的特性确定合适的杂交温度（T_{hyb}）（如果使用标准杂交液，靶序列 DNA 的 GC 含量为 40%，则 T_{hyb}为 42 ℃）。

② 计算杂交液的需要量，每 100cm² （10cm×10cm）的膜。预杂交需要 10ml 杂交液，杂交需要 3.5ml 杂交液，把预杂交液放在灭菌的塑料瓶中，在水浴中预热至杂交温度。

③ 把膜放在杂交袋中（注意：预杂交和杂交步骤可在任何密封的容器中进行，如导热性好的塑料或玻璃盒子、平皿、滚动的杂交瓶或密封的塑料袋，在杂交过程中，密闭性要好，防止孵育过程中 NH_3 的释放和 pH 的改变）。

④ 围绕着膜把杂交袋封边，尽量减少杂交袋的尺寸，从而降低预杂交液、杂交液和探针的用量。

⑤ 把预热的杂交液加入到杂交袋中（10 ml/100cm²）（注意：封边前要排除气泡，杂交袋中要有足量的杂交液，密封后看上去有点膨胀）。

⑥ 把杂交袋平放在摇动的水浴箱内，杂交袋的一角可坠上一重物，以保证完全浸入在水浴内。

⑦ 孵育至少 30min，在预杂交的过程中轻轻晃动膜，预杂交的时间可延长几小时，而不影响杂交结果。

5. Dig 标记的 DNA 探针的 Southern blot 杂交步骤

① 决定杂交中探针的需要量　如果检测人类基因组 DNA 中的单拷贝基因，推荐每毫升杂交液至少使用 25ng 随机引物标记探针杂交液或 2μl PCR 标记探针。

② 准备杂交液　从标记反应的管中取出适量标记探针；把探针加入一个无菌离心管中，再加 50μl 水，离心混匀；把无菌离心管放在沸水浴中 5min，变性探针；迅速插在冰浴上冷却探针；把变性的探针加到预热的杂交液中（3.5ml/100cm²），混匀。

③ 把杂交液加入到滤膜中　把密闭的杂交袋剪开，倒出预杂交液；立即加入含有 Dig 记探针的杂交液；排除气泡，把杂交袋封口。

④ 把杂交袋平放在摇动的水浴箱内，杂交袋的一角可坠一重物，以保证完全浸入在水浴内，在预定的杂交温度孵育 6～16h，孵育过程中不时轻轻摇动（建议杂交过夜，如果敏感性不要求非常高，孵育时间可以缩短。大多数杂交在 6h 即可以完成）。

⑤ 在塑料平皿中加 200ml 低严谨性缓冲液（2×SSC 含有 1g/L SDS）（注意：以下步骤中缓冲液的量取决于平皿的大小，必须保证膜能完成浸在液体中）。孵育结束后，剪开杂交袋，倒掉杂交液，把膜取出，立即放入盛有低严谨性缓冲液的塑料平皿中。

⑥ 室温，轻轻晃动平皿，孵育 5min，倒掉用过的缓冲液，立即加入 200ml 新鲜配制的低严谨性缓冲液，室温再孵育 5min，不时轻轻晃动平皿。

⑦ 预热高严谨性缓冲液，高严谨性缓冲液洗膜的温度依赖于探针和模板序列。例如，如果检测人或其他哺乳动物 DNA，参考表 4-3 进行调整。

⑧ 倒掉用过的低严谨性缓冲液，立即加入预热的高严谨性缓冲液洗膜，2×15min，温度为计算好的洗膜温度，不时轻轻摇动平皿。

为去除非特异性杂交，杂交后需要进行两步清洗。第一次清洗：低严谨性（高盐浓度和

表 4-3 探针和模板序列同源性对洗膜温度的影响

探针与靶序列的同源性/%	GC 含量/%	缓 冲 液	温度/℃
80～100	平均(40)	0.5×SSC+0.1%(体积分数)SDS	探针>100bp 为 65 探针≤100bp<65(据经验调整)
<80	平均(40)	0.5×SSC+0.1%(体积分数)SDS	大约 60(据经验调整)
80～100	高(50)	0.5×SSC+0.1%(体积分数)SDS	68

较低温度）去除非特异性结合的探针。第二次清洗：高严谨性（低盐浓度和较高温度）去除同源性较低的非特异杂交。

注：要预热高严谨性缓冲液，然后再清洗膜，否则短暂的冲洗不能破坏同源性较低的杂交混合物。

上述杂交过程中的 3 个重要影响因素是杂交的温度、盐离子浓度及甲酰胺浓度。一般规律是：温度高能增加杂交反应的特异性，降低温度则相反；高盐离子浓度降低反应的严谨性，低盐浓度则相反；甲酰胺能降低 DNA 的熔点，从而降低探针与靶基因的杂交温度，大约每增加 1%的甲酰胺，熔解温度降低 0.72℃。

二、Northern 印迹杂交法

Northern 印迹杂交的基本原理和过程与 Southern 印迹杂交基本相同，不同点在于：①所检测的靶核酸是 RNA，不是 DNA；②样品在电泳前不需要进行酶切；③样品在电泳前必须先进行变性；④为了防止 RNA 分子形成二级结构（发夹结构），维持其单链线性状态，整个电泳过程是在变性条件下进行的；⑤电泳结束后，不需要再进行变性和中和，可直接将 RNA 转移到膜上。以后的步骤与 Southern blot 基本相同。

1. RNA 样品的制备

制备 RNA 样品时应当注意防止 RNA 酶的污染，电泳分离时最好使用甲醛/MOPS 胶。操作过程可参照下列步骤进行。

（1）配制 1%～2%的琼脂糖凝胶（用于分离 RNA），用 1×MOPS 做缓冲液（注意：MOPS 溶液高压灭菌后变为淡黄色，但不影响使用），含 2%甲醛（例如，141.9ml 1×MOPS 缓冲液中，加 8.1ml 37%甲醛，溶解 1.5～3.0g 琼脂糖，制备凝胶）。

（2）1μg 总 RNA 或 100ng mRNA 上样，加 2 倍体积的新鲜配制的上样缓冲液，混匀。

（3）上样分子质量标准参照物。

（4）65℃ 10min 变性 RNA 和上样缓冲液的混合物。

（5）变性样品在冰上冷却 1min。

（6）如果上样缓冲液中含有甘油，则把样品加到上样孔中，然后加足量电泳缓冲液，进入步骤（7）。如果上样缓冲液中不含甘油，则进行下列步骤：①把样品加入到干胶中；②用电泳缓冲液小心充满上样孔；③加电泳缓冲液，刚及胶的顶部即可（不要没过加样孔）；④高电压（>3～4V/cm）电泳，大约 10min，样品则从加样孔迁移到凝胶中；⑤加足量电泳缓冲液，完全没过胶。

（7）3～4V/cm 电压进行电泳，使 RNA 很好地分离，至少 2 h 到过夜。

（8）为了评估 RNA 分离质量，电泳结束后，用 0.25～0.50μg/ml EB 染色，紫外灯下观察。

2. RNA 的转膜（虹吸印迹法）

对于 RNA 样品，转移前不建议应用碱处理，因为碱处理很容易破坏 RNA。应用虹吸印迹法转膜的实验步骤如下。

（1）在 20×SSC 中浸湿凝胶 15min，换液后重复浸胶 1 次，去除甲醛（甲醛抑制核酸的转移）。

（2）按下列操作步骤进行转膜，避免气泡的形成：①用 20×SSC 浸湿一张 Whatman 3MM 滤纸，放在支撑台上，形成一个“桥”；②把凝胶翻转放在浸湿的 Whatman 3MM 滤纸上，注意两者之间不要有气泡；③按照凝胶的大小剪一张尼龙膜；④把膜放在凝胶上；⑤在尼龙膜上放一张 Whatman 3MM 滤纸，一叠吸水纸，上置一玻璃板，其上压一个重约 200～500g 的物品。

（3）20×SSC 作为转移缓冲液（应无 RNA 酶），RNA 转移至少 6h，最好过夜。

（4）第二天，弃去膜以外的用品。

（5）当膜仍然湿润的时候，可用紫外交联法，把 RNA 固定在膜上：①把膜（DNA 面朝上）放在用 2×SSC 浸湿的 Whatman 3MM 滤纸上；②把湿膜暴露在紫外线下 1～3min；③在无菌双蒸水中浸湿膜；④空气中晾干膜。

也可用干烤法，在膜上固定 RNA：①在 2×SSC 中短暂洗膜；②120℃干烤固定 30min 或 80℃干烤固定 2h。

（6）如果不立即进行下一步，则可把膜保存在两张 Whatman 3MM 滤纸之间，放在密封袋中，4℃保存。

3. 预杂交

（1）杂交液需要量：通常使用杂交袋，每 100cm^2（10cm×10cm）的膜，预杂交时用 10～15ml 杂交液，杂交时用 3.5ml 杂交液，以保证在孵育的过程中有足量的缓冲液覆盖杂交膜。

（2）把上述计算的杂交液加入到灭菌离心管中，用水浴预热到 68℃。注意：对于大多数 RNA 与 RNA 的杂交，预杂交和杂交的温度是 68℃，然而，如果使用异源性的 RNA 探针，可以适当调整预杂交和杂交温度。

（3）把膜放到杂交袋中（注意：预杂交和杂交可在任何密封的容器中进行，如导热性好的塑料或玻璃盒子、平皿、滚动的杂交瓶或密封的塑料袋。在杂交过程中，密闭性要好，防止孵育过程中 NH_3 的释放和 pH 的改变）。

（4）加入适量预热杂交液（10～15ml/100cm^2 膜）。

注：要使用足够的杂交液，杂交袋封口后看上去有点膨胀。

（5）杂交袋平放在摇动的水浴箱内，在预杂交的过程中轻轻晃动膜，杂交袋的一角可坠上一重物，以保证完全浸没在水浴内，孵育至少 30min，预杂交的时间可延长几小时，而不影响杂交结果。

4. RNA 探针与 RNA 的杂交

Dig 标记的 RNA 探针杂交过程的实验步骤如下。

（1）确定杂交中探针的需要量　①推荐每毫升杂交液 100ng 应用体外转录法标记的 RNA 探针；②从标记反应的管中取出适量标记探针；③把探针加入到灭菌离心管中，再加 50μl DEPC 水，离心混匀；④把灭菌离心管放在沸水浴中 5min，变性探针；⑤迅速插在冰浴上冷却探针；⑥把变性探针加到预热的杂交液中（3.5ml/100cm^2），混匀。

（2）把杂交液加入到滤膜中　①把密闭的杂交袋剪开，倒出预杂交液；②立即加入含有

Dig 标记的 RNA 探针的杂交液；③排除气泡，把杂交袋封口。

(3) 把杂交袋平放在摇动的水浴箱内，杂交袋的一角可坠一重物，以保证完全浸入在水浴内，在 68℃杂交温度下，孵育 6～16h，孵育过程中不时轻轻摇动（推荐杂交过夜，然而如果不需要敏感性非常高，孵育时间可以缩短，大多数杂交在 6h 可以完成）。

(4) 在塑料平皿中加入低严谨性缓冲液［2×SSC，含有 0.1%（体积分数）SDS］，孵育结束后，剪开杂交袋，倒掉杂交液，把膜取出，立即放入盛有低严谨性缓冲液的塑料平皿中。

(5) 室温，轻轻晃动平皿清洗 5min，倒掉用过的缓冲液，立即加入 200 ml 新鲜配制的低严谨性缓冲液，室温再清洗 5min，不时轻轻晃动平皿。

(6) 预热高严谨性缓冲液［0.1×SSC，含有 0.1%（体积分数）SDS］到 68℃。

(7) 倒掉用过的低严谨性缓冲液，立即加入预热的高严谨性缓冲液洗膜，68℃洗膜 15min，不时轻轻摇动，换液后重复洗膜 1 次（注意：如果探针与靶 RNA 序列的同源性＜80%，则此步的洗膜温度可适当降低）。

三、斑点及狭缝印迹杂交

斑点及狭缝印迹杂交可用于基因组中特定基因及其表达的定性及定量研究。与 Southern 和 Northern 印迹法相比，其优点是简单、快速，可在同一张膜上进行多个样品的检测，常以此法确定最佳探针浓度。此法可同时检测同一样品中多个基因状态；但其缺点是不能鉴别所检测核酸的分子质量，特异性不高，有一定比例的假阳性。斑点杂交装置和检测结果如图 4-4 所示。

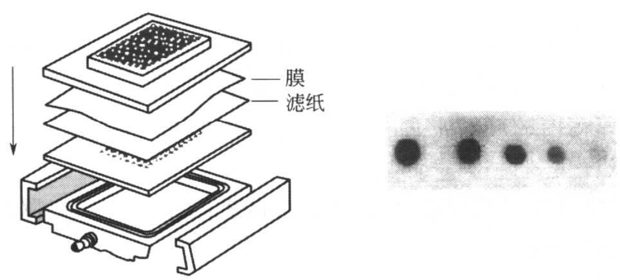


图 4-4　斑点杂交装置和检测结果示意图

为了将一定体积的核酸点到限定面积的滤膜上，一般采用真空点样器。点样器由多个可上样的孔组成，通过这些孔，在温和的抽真空的条件下，样品可流到膜上，成为一个个分开的点和条，然后用转移液冲洗点样孔，以保证样品全部流到滤膜上。将膜取下，烘烤或紫外照射固定核酸，再与特异探针进行杂交。定量研究时，可将待测核酸梯度稀释后分别点样。

四、菌落原位杂交法

首先，要获得生长在平板上的重组转化菌落，将其影印转移到膜上，经裂菌、DNA 变性、中和等处理后进行杂交，称为菌落原位杂交（colony hybridization *in situ*）。该法可用于基因工程研究中 DNA 重组体的筛选和临床标本的检测。

采用本方法对分散在若干个琼脂平板上的少数菌落（100～200）进行克隆筛选时，将这些菌落归并到一个琼脂主平板以及已置于第二个琼脂平板表面的一张硝酸纤维素滤膜上。经

培养一段时间后，对菌落进行原位裂解。主平板保存于4℃，直至得到筛选结果。

1. 将少数菌落转移到硝酸纤维素滤膜上

① 在含有选择性抗生素的琼脂平板上放一张硝酸纤维素滤膜。

② 用无菌牙签将各个菌落先转移至滤膜上，再转移至含有选择性抗生素但未放滤膜的琼脂主平板上。应按一定的格子进行划线接种（或打点）。每菌落应分别划线于两个平板的相同位置上。最后，在滤膜和主平板上同时划一个含有非重组质粒（如pBR322）的菌落。

③ 倒置平板，于37℃培养至划线的细菌菌落生长到0.5～1.0mm的宽度。

④ 用已装防水性黑色绘图墨水的注射器针头穿透滤膜直至琼脂，在3个以上的不对称位置作标记。在主平板大致相同的位置上也作上标记。

⑤ 用Parafilm封口膜封好主平板，倒置贮放于4℃，直至获得杂交反应的结果。

2. 菌落的裂解及DNA结合于硝酸纤维素滤膜

① 在一张保鲜膜上制作一个装有0.5mol/L NaOH的小凹陷（0.75ml），使菌落面朝上，将滤膜放到小凹陷上，展平保鲜膜，使滤膜均匀湿润，让滤膜留于原处2～3min。

② 用干纸巾从滤膜的下方吸干滤膜，用一张新的保鲜膜和新配制的0.5mol/L NaOH重复步骤①。

③ 吸干滤膜，将滤膜转移到新的带有1mol/L Tris-Cl（pH7.4）的保鲜膜凹陷上，5min后吸干滤膜，再重复一次该步骤。

④ 吸干滤膜，把它转移到有1.5mol/L NaCl、0.5mol/L Tris-Cl（pH7.4）的保鲜膜小凹陷上，5min后吸干滤膜，转移到一张干的滤纸上，置于室温20～30min，使滤膜干燥。

⑤ 将滤膜夹在两张干的滤纸之间，在真空烤箱中于80℃干烤2h，固定DNA。

3. 杂交

① 盛有2×SSC的塑料盘中，将烤干的滤膜飘浮在液面上，彻底浸湿5min。

② 将滤膜转到200ml预洗液的玻璃皿中。滤膜可叠在一起放于溶液中。用保鲜膜盖住玻璃皿，放到位于培养箱内的旋转平台上。于50℃处理30min。在这一步及以后的所有步骤中应缓缓摇动滤膜，防止它们粘在一起。

③ 用泡过预洗液的吸水棉纸轻轻地从膜表面拭去细菌碎片，以降低杂交背景而不影响阳性杂交信号的强度和清晰度。

④ 将滤膜转到盛有150ml预杂交液的玻璃中，在适宜温度（在水溶液中杂交时采用68℃，而在50%甲酰胺中杂交时采用42℃）下，预杂交1～2h。

⑤ 将^{32}P标记的双链DNA探针于100℃加热5min，迅速置于冰浴中。单链探针不必变性。将探针加到杂交袋中杂交过夜。杂交期间，盛滤膜的容器应盖严，以防液体蒸发。

⑥ 杂交结束后，去除杂交液，立即于室温把滤膜放入大体积（300～500ml）的2×SSC-0.1% SDS溶液中，轻轻振摇5min，并将滤膜至少翻转1次。重复洗1次，同时应避免膜干涸。

⑦ 68℃用300～500ml 1×SSC-0.1% SDS溶液洗膜2次，每次1～1.5h。此时已可进行放射自显影。如背景很高或实验要求严格的洗膜条件，可用300～500ml 0.2×SSC-0.1% SDS的溶液于68℃将滤膜浸泡60min。

⑧ 把滤膜放在纸巾上于室温晾干后，把滤膜（编号面朝上）放在一张保鲜膜上，并在保鲜膜上作几个不对称的标记，以使滤膜与放射性自显影片位置对应。

⑨ 用第二张保鲜膜盖住滤膜。加X光片并加上增感屏，于－70℃曝光12～16h。

⑩ 底片显影后，在底片上贴一张透明硬纸片。在纸上标记阳性杂交信号的位置，同时在不对称分布点的位置上作出标记。可从底片上取下透明纸，通过对比纸上的点与琼脂上的点来鉴定阳性菌落。

五、核酸原位杂交法

核酸探针的种类和标记物的不同，在具体应用的技术方法上也各有差异，但其基本方法和应用原则大致相同。大致可分为：①杂交前准备，包括固定、取材、载片和组织的处理，如何增强核酸探针的穿透性、减低背景染色等；②杂交；③杂交后处理；④显示，包括放射性自显影和非放射性标记的显色。

（一）组织与细胞的固定

原位杂交技术在固定剂的应用和选择上应兼顾三方面：①保持细胞结构；②最大限度地保持细胞内 DNA 或 RNA 的水平；③使探针易于进入细胞或组织。DNA 是比较稳定的，RNA 却非常容易被降解。因此，对于 DNA 的定位来说，固定剂的种类和浓度并不十分重要。相反，在 RNA 的定位上，如果要使 RNA 的降解减少到最低限度，那么不仅固定剂的种类浓度和固定的时间十分重要，而且取材后应尽快予以冷冻或固定。在解释原位杂交结果时应考虑到取材至进入固定剂或冰冻这段时间对 RNA 保存所带来的影响，因为组织中 mRNA 的降解很快。常用的固定剂有 10%甲醛、4%的多聚甲醛、乙醇-冰醋酸（3∶1，体积比）、戊二醛、Bouin 液、Carnoy 液等，其中最常用的是多聚甲醛，它不会与蛋白质产生广泛的交联，因而不会影响探针穿透入细胞或组织，这与戊二醛显著不同。乙醇-冰醋酸混合液和 Bouin 固定剂也能获得较满意的效果。对于 mRNA 的定位，常采用的方法是将组织固定于 4%多聚甲醛磷酸缓冲液中 1～2h，在冷冻前浸入 15%蔗糖溶液中，置 4℃冰箱过夜，次日切片或保存在液氮中。组织也可在取材后直接置入液氮冷冻，切片后才将其浸入 4%多聚甲醛约 10min，空气干燥后保存在－70℃。在－70℃可保存数月之久不会影响杂交结果。病理学活检取材多用福尔马林固定和石蜡包埋，这种标本对检测 DNA 和 mRNA 有时也可获得杂交信号，但石蜡包埋切片由于与蛋白质交联的增加，影响核酸探针的穿透，因而杂交信号常低于冰冻切片。同时，在包埋的过程会减低 mRNA 的含量。应用多聚甲醛蒸气固定干燥后的冷冻切片也可获得满意效果。各种固定剂均有各自优缺点，如沉淀性固定剂乙醇-冰醋酸混合液、Bouin 液、Carnoy 液等能为增加核酸探针的穿透性提供最佳条件，但它们不能最大限度地保存 RNA，而且对组织结构有损伤。戊二醛能较好地保存 RNA 和组织形态结构，但由于和蛋白质产生广泛的交联，从而大大地影响了核酸探针的穿透性。至今，多聚甲醛仍被公认为是原位杂交中较为理想的固定剂。

（二）载片和组织切片的处理

1. 载片的处理

玻片包括盖片和载片，载片的清洗至关重要，不能有任何核酸酶的污染。应采用热肥皂水刷洗，自来水清洗干净后，置于清洁液中浸泡 24h，清水洗净烘干，95%乙醇中浸泡 24h 后蒸馏水冲洗、烘干，烘箱温度最好在 160℃或以上，过夜烘干以去除任何 RNA 酶。盖玻片在有条件时最好用硅化处理，锡箔纸包裹无尘存放。

由于核酸原位杂交的实验周期长，实验程序繁杂，因此要应用黏附剂预先涂抹在玻片上，干燥后待切片时应用，以保证在整个实验过程中切片不致脱落。常用的黏附剂有铬矾-明胶液，其优点是价廉易得，但在长周期实验过程中，黏附效果不够理想。多聚赖氨酸液具

有较好的黏附效果，但价格昂贵。近年 Vector Lab（U. S. A.）推出一种新的黏附剂叫 Vectorband Reagent，每一单位包装可制备 500～700 张载玻片，黏附效果极佳，价格比多聚赖氨酸便宜，制片后可长期保存应用。

2. 增强组织的通透性和核酸探针的穿透性

此步骤根据应用固定剂的种类、组织的种类、切片的厚度和核酸探针的长度而定。如用戊二醛固定的组织，由于戊二醛与蛋白质产生广泛的交联，因此需要应用较强的增强组织通透性的试剂或方法，如应用稀释的酸洗涤、去垢剂 Triton X-100、乙醇，或某些消化酶如胃蛋白酶、胰蛋白酶、胶原蛋白酶和淀粉酶等。这种广泛的去蛋白作用无疑可增强组织的通透性和核酸探针的穿透性，提高杂交信号，但同时也会减低 RNA 的保存期及影响组织结构的形态，因此在用量及孵育时间上应慎为掌握。

在核酸原位杂交中，蛋白酶 K（Proteinase K）是应用于蛋白消化的关键酶，其浓度及孵育时间视组织种类、应用固定剂种类、切片的厚薄而定。一般用量是 1μg/ml（于 0. 1mol/L Tris-50mmol/L EDTA，pH8. 0 缓冲液中），37℃孵育 15～20min，以达到充分的蛋白消化作用而不致影响组织的形态为目的。蛋白酶 K 还具有消化包围着靶 DNA 的蛋白质的作用，从而提高杂交信号。在蛋白酶 K 消化后，应用 0. 1mol/L 甘氨酸溶液（在 PBS 中）清洗以终止蛋白酶 K 的消化作用，甘氨酸是蛋白酶 K 的抑制剂。为保持组织结构，通常用 4%多聚甲醛再固定。Burns 等（1987）报道应用胃蛋白酶（pepsin）20～100μg/ml（用 0. 1mol/L HCl 配）37℃、30min 进行消化，所获实验结果优于蛋白酶 K。

不少实验工作者在多聚甲醛固定后，浸入乙酸酐（acetic anhydride）和三乙醇胺（triethanolamine）中以减低静电效应，减少探针对组织的非特异性背景染色。但 Heinz、Hofer 等一些著名学者却对此持有异议，根据他们的实验和经验，乙酸酐和三乙醇胺液的处理并不能起到减低背景的目的，不能改善核酸原位杂交的信噪比。

3. 减低背景

核酸原位杂交实验程序中，如何减低背景染色是一个重要的问题。核酸原位杂交中背景染色的形成是诸多因素构成的。杂交后（Post-hybridization）的酶处理和杂交后的洗涤均有助于减低背景染色。

预杂交（prehybridization）是减低背景染色的一种有效手段。预杂交液和杂交液的区别在于前者不含探针和硫酸葡聚糖（dextran sulphate）。将组织切片浸入预杂交液中可达到封闭非特异性杂交点的目的，从而减低背景染色。

有的实验室在杂交后洗涤中采用低浓度的 RNA 酶溶液（20μg/ml）洗涤 1 次，以减低残留的和内源性的 RNA 酶，减低背景染色。

4. 防止 RNA 酶的污染

由于在手指皮肤及实验用玻璃器皿上均可能含有 RNA 酶，为防止其污染影响实验结果，在整个杂交前处理过程都需戴消毒手套。所有实验用玻璃器皿及镊子都应于实验前一日置高温（240℃）烘烤以达到消除 RNA 酶的目的。要破坏 RNA 酶，其最低温度必须在 160℃左右。有条件的实验室在消毒的玻璃器皿外包以锡箔纸以利于标记和防止取出时空气污染。在无高温消毒的烤箱时，也可用蒸汽消毒锅消毒。杂交前及杂交时所应用的溶液均需经高压消毒处理。

（三）杂交

核酸原位杂交的整个实验周期是比较长的，实验程序也比较繁杂，而杂交在整个实验中

可被认为是最直接接触的一步。杂交前的一切准备工作（如增加组织通透性）都是为了在杂交这一步骤中核酸探针能进入细胞或组织，与其内的靶核苷酸相结合。因此，杂交是核酸原位杂交中关键的而且是最重要的一个环节。

杂交是将杂交液滴于切片组织上，加盖硅化的盖玻片（或采用无菌的蜡膜代替硅化的盖玻片）可获得满意的实验结果。加盖片的目的是防止孵育过程中的高温（50℃左右）导致杂交液的蒸发。因此，也有为稳妥起见，在盖玻片周围加液体石蜡封固的，但有研究者认为这并不十分必要，因封固的石蜡在高温下融解反易导致杂交液的污染，必要时可加橡皮泥封固盖片四周。硅化的盖玻片的优点是清洁，无杂质；光滑，不会产生气泡和影响组织切片与杂交液的接触；盖玻片自身有一定重量，能与有限的杂交液吸附，达到覆盖和防止蒸发的作用。在孵育时间较长时，为保证杂交所需的湿润环境，可将带有硅化盖玻片进行杂交的载玻片放在盛有少量 5×SSC 或 2×SSC 溶液的硬塑料盒（要能防止高温破坏）中进行孵育。杂交液的成分和预杂交液基本相同，所不同的是加入了标记的核酸探针和硫酸葡聚糖。

为了获得满意的实验结果，在杂交实验中还须注意以下的环节。

（1）探针的浓度　很难事先确定每一种实验探针的浓度，但要掌握一个原则，即探针浓度必须给予该实验最大的信噪比。背景染色的高低也与探针浓度有关。根据国内外实验工作者的经验，认为最佳原则应是应用最低探针浓度以达到与靶核酸的最大饱和结合度为目的。这与在免疫细胞化学实验中选择抗血清的最佳工作浓度的原则相同。

探针浓度依其种类和实验需要略有不同，在原位杂交细胞化学中，探针浓度为 0.5～5.0μg/ml（即 0.5～5.0ng/μl）。根据 Heinz 、Hofelt 实验室经验，对放射性标记的 dsDNA 或 cRNA 探针，其浓度在 2～5ng/μl。Conlton 认为生物素标记的探针，其最佳浓度在0.5～5ng/μl。

必须强调的是，国内外实验室都证明加杂交液的量要适当，以每张切片 10～20μl 为宜。杂交液过多不仅造成浪费，而且液量过多常易致盖玻片滑动脱落，影响杂交效果，过量的杂交液含核酸探针浓度过高，易导致高背景染色等不良后果。

（2）探针的长度　一般应用于核酸原位杂交探针的最佳长度应在 50～100 个碱基之间。探针短则易进入细胞，杂交率高，杂交时间短。据报道，长 500 个碱基的探针，其杂交时间约需 20h。200～500 个碱基的探针仍可应用，超过 500 个碱基的探针在杂交前最好用碱或水解酶进行水解，使其变成短的片段，达到实验所要求的碱基数。

（3）杂交的温度和时间　杂交的温度也是杂交成功与否的一个重要环节。DNA 或 RNA 需变性解链后才能进行杂交。原位杂交中，多数 DNA 探针需要的 T_m 是 90℃，而 RNA 则需要 95℃。这种高温对保存组织形态完整和保持组织切片黏附在载玻片上是不可能的。因此，在杂交的程序中常规性的加入 30%～50%甲酰胺（formamide）于杂交液中。McConaughy 报道，反应液中每增加 1%的甲酰胺浓度，T_m 值可降低 0.72℃。因此，可用调节盐浓度的办法来调节 T_m。由于盐和甲酰胺浓度的调节等因素，实际采用的原位杂交的温度在 T_m－25℃左右，即比 T_m 减低 25℃，大约在 30～60℃之间，根据探针的种类不同，温度略有差异，RNA 和 cRNA 探针一般在 37～42℃左右，而 DNA 探针或细胞内靶核酸为 DNA 时，则必须在 80～95℃加热使其变性，时间 5～15min（有报道在 105℃微波炉加热使之变性），然后在冰上放置 1min，使之迅速冷却，以防复性，再置入盛有 2×SSC 的温盒内，在 37～42℃孵育杂交过夜。

杂交的时间过短会造成杂交不完全，而过长则会增加非特异性杂交。从理论上讲，核酸

杂交的有效反应时间在3h左右。Barns等（1987）报道用DNA探针杂交，其反应完成时间为2～4h。但为稳妥起见，一般将杂交反应时间定为16～20h，或为简便起见杂交孵育过夜。当然，杂交反应的时间与核酸探针长度和组织通透性有关，在确定杂交反应时间应予考虑，并经反复实验确定。

有作者主张杂交反应的孵育应在黑暗环境中进行，因为光线会促进甲酰胺的电离作用。

（4）杂交严谨度（hybridization stringency） 杂交条件的严谨度（stringency）表示通过杂交及冲洗条件的选择对完全配对及不完全配对杂交体的鉴别程度。错配对（mismatch）杂交的稳定性较完全配对杂交体差，因此，通过控制杂交温度、盐浓度等，可减弱非特异性杂交体的形成，提高杂交的特异性。所以，杂交的条件愈高，特异性愈强，但敏感性降低，反应亦然。一般来说，低严谨度（low stringency）杂交及冲洗条件在 $T_m-35℃$ 至 $T_m-40℃$ 之间，高盐或低甲酰胺浓度。在这种条件下，大约有70%～90%的同源性核苷酸序列被结合，其结果是导致非特异性杂交信号的产生。中严格度的杂交温度为 $T_m-20℃$ 至 $T_m-30℃$ 的范围。高严谨度（high stringency）为 $T_m-10℃$ 至 $T_m-15℃$，低盐和高甲酰胺浓度。在这种条件下，只有具有高同源性的核苷酸序列才能形成稳定的结合。

由于原位杂交技术多数是在 $T_m-25℃$ 进行的，不属于高严谨范围，无疑会产生非特异性结合，导致信噪比降低。在这种情况下，可加强杂交后处理洗涤的严谨度，使非特异性的杂交体减少。由于RNA杂交的稳定性，应用cRNA探针进行细胞或组织的原位杂交时的杂交温度比其他核酸探针要高10～15℃。实验证明，cRNA产生的信号比双链cDNA要强。单链的RNA探针其杂交信号大于双链的cDNA的约8倍。

（5）硫酸葡聚糖（dextran sulphate）和甲酰胺（formamide） 在杂交液中，甲酰胺占50%左右，硫酸葡聚糖占10%左右。硫酸葡聚糖是核酸杂交液中仅次于甲酰胺的一种组成成分。它是一种大分子的多聚胺化合物，具有极强的水合（hydrate）作用，因而能大大增加杂交液的黏稠度。硫酸葡聚糖的主要作用是促进杂交率，特别是对双链核酸探针。

甲酰胺在调节杂交反应温度方面起了极为重要的作用，它有助于保持组织的形态结构。甲酰胺还可防止在低温时非同源性片段的结合，但甲酰胺可破坏氢键，从而具有一种不稳定的作用。

（四）杂交后处理

杂交后处理（post-hybridization treatment）包括系列不同浓度、不同温度的盐溶液的漂洗。在原位杂交的实验程序中，这也是一个重要的环节。大多数的原位杂交实验是在低严谨度条件下进行的，非特异性的探针片段黏附在组织切片上，从而增强了背景染色。RNA探针杂交时产生的背景染色特别高，但能通过杂交后的洗涤可有效地减低背景染色，获得较好的反差效果。在杂交后漂洗中的RNA酶液洗涤能将组织切片中非碱基配对RNA除去。洗涤的条件，如盐溶液的浓度、温度、洗涤次数和时间因核酸探针的类型和标记的种类不同而略有差异，一般遵循的共同原则是：洗涤时，盐溶液浓度由高到低，而温度由低到高。必须注意，在漂洗的过程中，切勿使切片干燥。干燥的切片即使用大量的溶液漂洗也很难减少非特异性结合，从而增强了背景染色。放射性标记探针杂交后漂洗过程中可用底片曝光的方法检测背景（非特异性标记的多少）作为改善漂洗程序的指针。^{35}S 标记的核酸探针在漂洗液中须加入14mmol/L的β-巯基乙醇（β-mercaptoethanol）或硫代硫酸盐（thiosulphate），以防止 ^{35}S 标记的核酸探针被氧化。总之，如何控制漂洗的严谨度从而达到理想的信噪比无既定方案可循，必须从反复的实践中取得经验。

（五）显示

显示又称为检测系统（detection system）。根据核酸探针标记物的种类分别进行放射自显影或利用酶检测系统进行不同显色处理。

细胞或组织的原位杂交切片在显示后均可进行半定量的测定，如放射自显影可利用人工或计算机辅助的图像分析检测仪检测银粒的数量和分布的差异。非放射性核酸探针杂交的细胞或组织可利用酶检测系统显色，然后利用显微分光光度计或图像分析仪对不同类型和数量的核酸的显色强度进行检测。但利用核酸原位杂交做半定量测定必须注意严格控制实验的同一条件、切片的厚度和核酸的保存量如取材固定的间隔时间等；如为放射自显影，核乳胶膜的厚度和漂洗液的稀释度等必须保持前后一致。

（六）对照实验和核酸原位杂交结果的判断

核酸原位杂交实验中，并非任何阳性信号都是特异性的，故必须同时有对照实验以证明其特异性。阳性对照可选择：①Northern 或 Southern 印迹杂交；②核酸原位杂交与免疫细胞化学结合对照；③应用多种不同的核酸探针与同一靶核酸进行杂交对照。阴性对照可选择：①将 cDNA 或 cRNA 探针进行预杂交（吸收试验）对照；②与非特异性（载体）序列和不相关探针杂交（置换试验）对照；③将切片应用 RNA 酶或 DNA 酶进行预处理后杂交对照。此外，还有用同义 RNA 探针（sense probe）进行杂交的对照、不加核酸探针杂交液进行杂交（空白试验）对照、用已知确定为阳性或阴性组织进行核酸原位杂交对照、用未标记探针做核酸原位杂交对照等方法。

从理论上讲，对照试验设置愈多，特异性确定愈可靠，但实际对照是有限的。因此，应任选设至少 3～4 种用以证实核酸原位杂交结果的可靠性。在上述实验中，①Northern 和 Southern 印迹杂交法证明的方式和用 Western 印迹法检测抗体（蛋白质）的特异性一样，是比较可靠的。②如果具备相当的免疫组化抗血清，可用结合的免疫组织化学和核酸原位杂交法从蛋白质（或多肽）水平和转录水平在相邻切片或同一切片中证明同一种多肽和相应 mRNA 共存于同一细胞中。③预先将切片用 DNA 酶或 RNA 酶消化，然后用核酸原位杂交技术证明丢失的是 DNA 或 RNA。如同免疫组化的吸收实验一样，事先与特异性的 cRNA 或 cDNA 进行杂交，再进行核酸原位杂交，其结果应为阴性。④由于同义 RNA 探针和组织内 mRNA 序列顺序是相同的，应用其进行核酸原位杂交，结果应为阴性。检测系统的对照如乳胶或酶显色系统也应在无标记探针的情况下进行。

核酸原位杂交的最大优点是它的高度特异性，它可测定组织、培养的单个细胞或细胞提取物中的核酸含量。应用高敏感度的放射性标记 cRNA 探针在理想的核酸原位杂交的实验条件下检测 mRNA，其敏感度可达到每个细胞 20 个 mRNA 拷贝。由于双链 DNA 的稳定性，在用核酸原位杂交定位 DNA 时很少发生丢失、降解。在靶核酸序列比较伸展的情况如染色体铺片，长于 2kb 的探针可以应用。因此，其敏感性高到能够出在染色体铺片上，有时甚至在组织切片上的单个基因拷贝。正因为如此，对核酸原位杂交结果的解释应持慎重态度，特别是前人未报道过的新发现。因为如前所述，影响核酸原位杂交实验结果的因素太多，比如在外科或实验取材后未及时固定或冷冻，可由于组织中 mRNA 的降解而导致假阴性结果。另外，各种类型核酸探针进入细胞、组织和各种器官的能力（又叫可接近性，acessiblity）各异。这些因素都将影响核酸原位杂交的实验结果。

综上所述，核酸原位杂交中将组织或细胞固定在载片上，采用去垢剂和（或）蛋白酶对组织细胞内的蛋白质进行部分性地消化，使探针在细胞基质中获得最大的穿透力，应用

DNA 探针、RNA 探针、寡核苷酸探针等，采用同位素标记或非同位素标记。若应用的是放射性核素标记探针，杂交后必须经过放射自显影，以对被检测的核酸进行定位与定量；若所用的是非放射性标记的探针，则根据标记体系对杂交结果进行检测，如生物素标记的探针主要用抗生物素（biotin）抗体和亲和素（avidin）/链霉亲和素（streptavidin）检测。

原位杂交技术为核酸序列在细胞水平的定位与测定提供了直接的方法。它可查明染色体中特定基因的位置，为染色体疾病的诊断和研究提供了方法；也可通过 cDNA 探针检测 mRNA 是否存在，来反映细胞内特定基因的表达。原位杂交的结果可以显示有关核酸序列的空间位置状况及含该核酸序列的具体细胞，包括细胞的具体定位、数目及类型。

六、Western 印迹杂交法

Western 印迹杂交法自 1979 年由 Towbin 等引入检测蛋白质以来，已经成为一项常用的技术。它结合了凝胶电泳的高分辨率和固相免疫测定的高特异性，被广泛用于蛋白质研究。Western 印迹杂交法技术无需对靶蛋白进行放射性核素标记，因此是一种十分安全有效的检测方法。Western 印迹杂交法的大致步骤是：凝胶电泳分离蛋白质，将电泳结果电转移至合适的膜上，标记的抗体或抗原与靶蛋白质的结合，染色观察或直接鉴定。

（一）蛋白质的分离

不同样品的蛋白质分离方法有所不同，可参照各种常规蛋白质制备手册或方法。一般 Western 印迹检测蛋白质的灵敏度为 1～5ng，普通厚度的 SDS-聚丙烯酰胺凝胶（SDS-PAGE）单孔中蛋白质的加样量约为 100μg，因此样品中靶蛋白占总蛋白的比例为 $1/10^5$，即可被检测出来。蛋白质在 SDS-PAGE 中的分离主要依据其分子质量的差异，此时蛋白质处于变性状态。而有时需要分离检测天然状态的蛋白质，这就需要进行非变性 PAGE 电泳分离。更为精细的分离需要借助二维电泳，即等电聚焦电泳和 PAGE 联用，这种分离方法在目前的蛋白质组研究中广泛采用。对普通的培养细胞样品进行 SDS-PAGE 分离蛋白质，常采用下列方法。

① 细胞经 PBS 溶液洗 1～2 次后用 SDS 加样缓冲液溶解，用细胞刮子刮取细胞，溶液转移至微量离心管中。

② 沸水浴中煮 5～10min。

③ 超声波处理打碎染色体 DNA，直到溶液不黏为止。

④ 室温，10000r/min 离心 10min，取上清，转移至另一离心管中。

⑤ 测定蛋白质浓度，进行 SDS-PAGE。

电泳中 PAGE 胶的浓度通常为 10%，它适合分离分子量在 10000～130000 的蛋白质分子。如果是未知蛋白质或蛋白质分子量范围很宽，则推荐使用梯度胶。普通的 SDS-PAGE 胶的厚度是 1mm 或 1.5mm 厚，而在 Western 印迹中使用的胶厚度应小于 1mm，以有利于转膜。

（二）转膜

Western 印迹的转膜方法同样有虹吸法、电转移法、真空转移法三种，其中电转移法由于速度快、转移完全而成为最普遍采用的方法。转移采用的膜主要有硝酸纤维素滤膜和聚偏氟乙烯膜（polyvinylidene fluoride，PVDF），前者价格便宜而性能稍差，后者正好相反。实验中可根据具体情况选用。

电转移方法有 2 种转移系统：槽转移系统和半干转移系统，如图 4-5 所示。

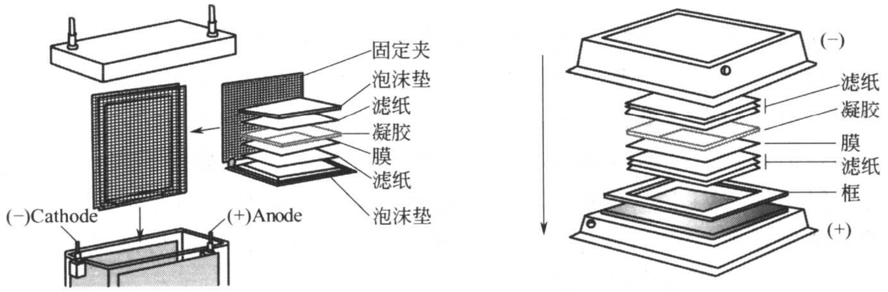


图 4-5　槽转移系统和半干转移系统示意图

上述 2 个转移系统的原理相同，但操作上槽转移系统更为方便和常用。转移系统的操作并不复杂，具体过程可参照产品使用说明进行。下面以槽转移系统为例说明有关转移参数。通常，影响转移的因素有转移缓冲液、电流和转移时间、平衡时间等。

1. 转移缓冲液

缓冲液提供一个稳定的电流环境，同时保证样品中的蛋白质不会沉淀析出。通常使用的缓冲液是 pH8.3 的 Tris-甘氨酸缓冲液，这一 pH 高于大多数蛋白质的等电点，因此，蛋白质在此缓冲液中带负电荷，朝阳极移动。槽转移系统中由于缓冲液均匀充满槽中，因此转移过程中带电离子分布相对恒定。转移缓冲液中常需加甲醇，其主要作用是稳定胶的尺寸、从蛋白质中剥离 SDS，这是因为电泳凝胶在水性的缓冲液中会膨胀，而加入甲醇能减小这一作用。甲醇对 SDS 的剥离作用对蛋白质转移至膜上至关重要，尽管 SDS 对蛋白质的电泳分离很重要，但对蛋白质转移至膜上有阻碍，加入甲醇后这一阻碍消除。

2. 电流和转移时间

合适的转移电流与时间对充分有效的蛋白质转移非常重要。电流强度太高可能导致蛋白质移动太快而未能吸附在膜上，这对小分子蛋白质尤其突出。因此应参照使用手册，控制合适电流大小。也可根据实验的具体情况（如胶的浓度、缓冲液的组成、分离蛋白质分子质量等）优化转移条件。一般对槽转移系统应采用较长的转移时间，并加上内循环冷却池辅助转移。

3. 平衡时间

早期的印迹实验在转移前需要将凝胶在缓冲液中预先平衡较长时间，通常为 30min。现在由于更多采用小胶（迷你胶，mini-gel）转移，因此平衡时间大为缩短，只需 5～10min 胶的大小即不再变化。然而由于 SDS 的剥离速度相对较慢，因此实际需要的平衡时间应适当延长，一般推荐最少 15min 的平衡时间。但应注意，对于小肽，由于其在缓冲液移动扩散速度快，因此平衡时间应限制在 10min 以内，以防止蛋白质扩散损失。

（三）鉴定用膜的准备

1. 膜的干燥

转移完成以后，应该在检测鉴定前完全使膜干燥。这一过程将有利于蛋白质牢固结合在膜上。随着膜的干燥，膜会变得越来越不透明，这种肉眼所见的变化常会掩饰膜内微孔中所含有的水分，应继续完全干燥微孔中的水分，这样干燥的膜在−70℃可长期存放，在−20℃也可存放 2 个月时间，而不影响今后的检测。干燥的膜在重新使用之前应在 100%甲醇中浸泡后，才能使用。

2. 显色

显色技术相对简单，它使膜上的蛋白质条带能够显现出来。显色包括2种方法：一是染色；二是透明显色。显色的主要目的是观察转移的效果以及各泳道中蛋白质的均衡情况。染色方法可参照一般蛋白质的染色方法，包括可逆的和不可逆的染色方法。一般不可逆的染色方法灵敏度高，但影响后续的检测鉴定；可逆的方法则相反。通常考虑可逆的染色方法。另外，对于有些膜（如PVDF）可采用透明显色的方法观察转移的情况。在PVDF膜上，有蛋白质条带的位置能被20%的甲醇浸湿而呈现透明的区域，这样即可观察到膜上蛋白质的分布情况，这种方法显然简单而方便。

（四）鉴定

Western印迹杂交的鉴定有多种方法，其中最常用的是免疫学鉴定方法。免疫检测使用特异性的抗体，结合并定位膜上的斑点（印迹），因此能在复杂的蛋白质混合物中鉴定出特异的靶蛋白质分子。标准的免疫检测过程包括如下步骤：

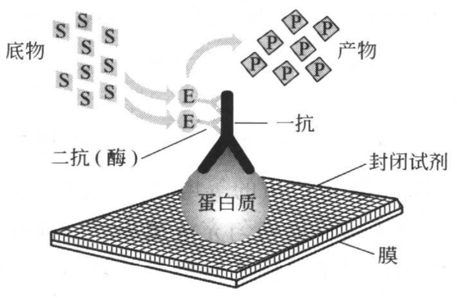


图4-6　膜基础的免疫检测示意图

① 阻断膜上未结合的区域，防止抗体的非特异性结合；

② 用一抗与膜孵育，使其结合目标蛋白；

③ 洗膜移去未结合的一抗；

④ 用能识别一抗的交联二抗与膜孵育；

⑤ 洗膜移去未结合的二抗；

⑥ 用能与交联二抗反应的底物与膜孵育，显示特定蛋白的位置。

上述过程如图4-6所示。

（刘志国　王淳本）

第五章 分子克隆技术

“克隆”（clone）一词目前被广泛使用，其含义是指从一个共同祖先无性繁殖下来的一群遗传上同一的DNA分子、细胞或个体所组成的特殊的群体；或指产生这一群体的过程。克隆技术迅速发展，目前已经成为分子生物学领域重要的基本技术之一。分子克隆（molecular cloning）也称基因克隆，是指在体外，借助于能自我复制的载体分子，将目的基因引入宿主细胞中进行增殖，以获得大量单一的特定DNA片段，或以此为基础，通过工程方法表达出大量的相应蛋白质产物。

本章将重点介绍分子克隆技术的基本过程及密切相关的重要技术，包括克隆中需要的工具酶、克隆载体、目的基因的获得及克隆筛选鉴定技术。

第一节 概 述

分子克隆操作的主要对象是基因。待研究或待检测的特定基因被称为目的基因（target gene）。目的基因通过相应的基因载体导入宿主细胞，并借助宿主的细胞体系进行复制或表达，所以对宿主细胞DNA而言，目的基因被称为外源基因或外源DNA。分子克隆的目的就是研究基因与改造基因。根据研究目的的不同，研究对象有所不同，目的基因可大致分为三类：病毒基因、原核基因及真核基因。病毒基因组与原核生物基因组相对简单，较易获得目的基因；而真核生物基因组庞大复杂，要从众多的基因中获得所需的目的基因难度相对较大。

一、分子克隆的基本步骤

分子克隆技术目前已经有了很大的发展，取得了巨大的成就，特别是一些新技术和先进研究设备的不断涌现，进一步推动了分子克隆技术的发展；但其基本研究步骤仍主要包括以下几个方面。

① 采用各种方法从复杂的生物体基因组中分离获得带有目的基因的DNA片段。

② 在体外，将带有目的基因的外源DNA片段连接到具有自我复制功能及筛选标记的载体分子上，构建成重组DNA分子。

③ 将重组体转移到宿主细胞，并随宿主细胞的繁殖而扩增。

④ 从细胞繁殖群体中筛选出获得了重组体的受体细胞克隆。

⑤ 从筛选出来的受体细胞中提取已经得到扩增的目的基因，以做进一步的分析鉴定。

⑥ 将目的基因克隆到合适的表达载体上，导入宿主细胞，构建成高效、稳定的具有功能性表达能力的基因工程细胞。

⑦ 利用工程技术大规模培养上述基因工程细胞，获得大量的外源基因表达产物。

⑧ 工程细胞表达产物的分离纯化，并最后获得所需的基因工程产品。

上述8个步骤也可归并为两大部分，分属上游技术和下游技术。其中上游技术包括①～⑤；下游技术包括⑥～⑧。两大部分有机结合成为一个整体：上游技术是分子克隆的核心与基础，下游技术则是克隆基因的应用。上、下游技术紧密衔接，相互兼顾，将有力地促进克隆技术的发展与完善。

二、分子克隆研究的内容

分子克隆研究的主要内容包括以下几个方面。

（一）分子克隆工具的研究

分子克隆技术之所以能够在体外将不同来源的DNA重新组合构成新的DNA分子并在宿主细胞扩增和表达，主要依赖一系列重要的克隆工具，主要包括：分子克隆载体、分子克隆的工具酶及受体系统。对分子克隆技术中这些重要的工具的研究是分子克隆研究的重要内容之一。载体是外源基因的运载工具，它能带动外源基因在宿主细胞内复制和表达，并能利用载体上特殊的筛选标记获得含有重组体的细胞克隆，因此是基因克隆操作中必备的工具。目前已构建了数以千计的各类载体，其研究与应用极大地推动了基因克隆工作的发展。工具酶是指体外进行DNA切割、连接、修饰及合成等过程中所需要的酶，主要包括：限制性核酸内切酶、DNA连接酶、DNA聚合酶及各种修饰酶等。例如，限制性核酸内切酶能在特定位点切割DNA，连接酶能使不同来源的DNA片段相互连接。因此，工具酶是分子克隆操作的基础。受体是载体的宿主，是外源基因表达的场所；它可以是单个细胞，也可以是组织、器官，甚至是个体。受体的选择需要根据实验的目的、所用的载体以及所采用的实验方法而定。目前最为常用的受体系统是大肠杆菌受体系统和酵母受体系统，它们分别是原核与真核受体的代表，除此以外，也发展了其他受体系统，如链霉菌、芽孢杆菌、丝状真菌及动物细胞受体系统，还有的利用动物的组织（如乳腺组织、胚胎组织等）作为受体进行基因表达的研究等，由此可见受体系统的研究也是克隆研究中的一项重要内容。

（二）分子克隆技术的研究

分子克隆诞生至今只有短短30年余的时间，但这项技术已经取得了巨大的发展，并广泛渗透到各个学科，推动着整个生命科学的发展。在这期间基因克隆技术的发展起到了重要的作用，新的克隆方法不断涌现，如以PCR为基础的差异筛选技术、长片段DNA序列测定技术、高通量的基因芯片技术、借助计算机和互联网的生物信息技术等，这些研究技术的运用彻底改变了分子克隆研究的规模和速度，成为分子克隆研究中的重要内容。

（三）克隆对象——目的基因的研究

基因是一种重要的生物资源，而且是一种有限的战略性资源，目前各国都十分重视基因资源的开发，倾力资助基因研究。根据人类基因组计划的最新研究结果，人类基因总数比预期的低得多，初步定位在2.6万～4.0万个之间，如此少的基因自然成为激烈竞争的目标，谁拥有的基因专利多，谁就在分子克隆领域占据了主动地位。由此可见，获得目的基因也同样是分子克隆研究的重要内容，这方面的研究工作已经从零星的单个基因的研究发展到大规模的基因组研究；从人类基因组延伸其他生物基因组。

（四）分子克隆产品的研究

研究基因除了分析基因的结构与功能以外，更具魅力的是研究基因的表达产物及其在各

领域里的应用，其中最突出的是在医药领域的应用。从 1982 年美国 Lilly 公司首先将重组胰岛素投放市场以来，迄今已有 50 多种基因工程药物上市，近千种处于研发状态，形成一个巨大的高新技术产业，产生了不可估量的社会效益和经济效益。据不完全统计，美国的生物技术公司已有 2000 多家，欧洲共有约 1000 家，其中以英、德、法为主。正在开发的生物药物按分子类别归类，排在前九位的分别是单克隆抗体、疫苗、基因治疗、白细胞介素、干扰素、生长因子、重组可溶性受体、反义药物和人生长激素。目前基因工程药物的发展正在带来一场新的药业革命，它将成为 21 世纪药业的支柱。

第二节 克隆操作中需要的工具酶

分子克隆技术中，DNA 分子的切割、重组、修饰及合成都涉及一系列酶促反应。催化这些反应的酶是进行 DNA 操作的基本工具，故称为工具酶。用于分子克隆的工具酶种类繁多、功能各异，如限制性核酸内切酶的内切功能、DNA 连接酶的连接功能、末端脱氧核糖核酸转移酶的核苷酸转移功能等，它们在基因的切割、连接及修饰等诸多方面发挥着重要作用。此外还有 DNA 聚合酶、逆转录酶和碱性磷酸酶等也都是基因克隆技术中常用的工具酶。正是由于这些工具酶的发现和应用，才使得 DNA 的体外重组得以实现，并促进了分子克隆技术的不断发展。表 5-1 列举了克隆操作中几种常见的工具酶及其主要用途。

表 5-1 基因克隆操作中常用的工具酶

酶	主 要 用 途
限制性核酸内切酶	识别 DNA 特定序列，切割 DNA 分子
DNA 连接酶	连接两个 DNA 分子或片段
大肠杆菌 DNA 聚合酶Ⅰ或其大片段(Klenow 酶)	①缺口平移制作标记 DNA 探针；②合成 cDNA 的第二链；③填补双链 DNA 3′-隐蔽末端；④DNA 序列分析
Taq DNA 聚合酶等	聚合酶链反应(PCR)
多核苷酸激酶	催化多核苷酸 5′-羟基末端磷酸化，制备末端标记探针
末端转移酶	在 3′-末端加入同聚物尾
S1 核酸酶	降解单链 DNA 或 RNA，使双链 DNA 突出端变为平端
DNase Ⅰ	降解 DNA，在双链 DNA 上产生随机切口
RNase A	降解 RNA
磷酸酶	切除核酸末端磷酸基

一、限制性核酸内切酶

通过切割相邻的两个核苷酸残基之间的磷酸二酯键，导致核酸分子多核苷酸链水解断裂的酶叫做核酸酶。其中专门水解 RNA 分子的核酸酶叫做核糖核酸酶（RNase），而特异水解 DNA 分子的则叫做脱氧核糖核酸酶（DNase）。核酸酶按其水解核酸分子的不同方式，又可

分为两种类型：一类是从核酸分子的末端开始，逐个降解核苷酸，这种酶称为核酸外切酶（exonuclease）；另一类是从核酸分子的内部水解磷酸二酯键，使之断裂成小片段，称为核酸内切酶（endonuclease）。在分子克隆的操作过程中，核酸内切酶具有特别重要的意义。

限制性核酸内切酶（简称限制酶）是一类能识别双链 DNA 分子中特定核苷酸序列，并在识别序列内或附近特异切割双链 DNA 的核酸内切酶。限制酶天然存在于细菌体内，与相伴存在的甲基化酶共同构成细菌的限制修饰体系，称限制-修饰系统，它是原核生物中普遍存在的一种防御系统，对维持细菌遗传物质的稳定有重要意义，目前使用的限制酶都是从原核生物中分离获得。

（一）限制性核酸内切酶的类别

到目前为止，已从数百种不同的微生物中，发现了大约上千种能切割特异 DNA 序列的限制酶，同时发现有百种以上不同的切割位点。根据限制酶的作用特点将其分为三型，即Ⅰ型、Ⅱ型和Ⅲ型，并按 H. Smith 和 D. Nathams 提出的命名原则对其系统命名。

1. Ⅰ型与Ⅲ型限制性核酸内切酶

Ⅰ型酶是最早分离的限制酶，它属复合功能酶，兼具限制、修饰两种功能。需要 Mg^{2+}、ATP 和 *S*-腺苷甲硫氨酸作为催化反应的辅助因子，在降解 DNA 时，伴有 ATP 的水解。因此它具有核酸内切酶、甲基化酶、ATP 酶和 DNA 解旋酶 4 种活性。其显著特点在于酶的识别位点与切割位点不一致，即没有固定的切割位点，切割位点一般位于识别位点 400bp 以外（最多可达 7000 kb），并随机切割 DNA，不产生特异性 DNA 片段，所以在基因重组中没有应用价值。Ⅰ型酶的典型代表 *Eco*K 和 *Eco*B，它们分别来致大肠杆菌 K 株和 B 株，两者的分子量相似，均为 300000。它们均由三种亚基组成，其中，特异性亚基（即 γ 多肽链）具有特异性识别 DNA 序列的活性，修饰亚基（即 β 多肽链）具有甲基化酶的活性，限制亚基（α 多肽链）具有核酸内切酶活性；不同的是，*Eco*B 的修饰酶可与其限制酶分开，即 α 亚基可与 β、γ 亚基分开。

Ⅲ型酶也有核酸内切酶和甲基化酶作用，酶分子由两个亚基组成，其中 M 亚基（修饰亚基）负责位点识别与修饰，R 亚基（限制亚基）则具有核酸酶活性。Ⅲ型酶在 DNA 链上有特异的切割位点，其切割位点在识别位点以外，距识别位点 3′端 24～26bp 处，切割反应需要 ATP、Mg^{2+} 和 *S*-腺苷甲硫氨酸的激活。目前知道的Ⅲ型酶数量很少，在分子克隆中实际作用不大。

2. Ⅱ型限制性核酸内切酶

Ⅱ型酶就是通常所指的 DNA 限制性核酸内切酶。Ⅱ型酶限制-修饰系统分别由限制酶与修饰酶两种不同的酶所组成。Ⅱ型限制酶分子量小，仅需 Mg^{2+} 作为催化反应的辅助因子，它们能识别双链 DNA 的特殊序列，并在这个顺序内进行切割，产生特异的 DNA 片段。Ⅱ型限制性核酸内切酶种类多，并且可以特异地切割 DNA，产生特异性的片段，是克隆技术中最重要的内切酶。

绝大多数的Ⅱ型限制性核酸内切酶可以识别长度为 4～8 个核苷酸的特异性序列，称为核酸内切限制酶识别序列。这些序列中，有些是连续的（如 GATC），有些则是间断的（如 GANTC）。许多识别序列是具有 180°旋转对称的反向重复序列，也称回文结构，即同一条单链以中心轴对折可形成互补的双链，如 *Hind*Ⅲ 和 *Bam*HⅠ 的识别位点分别为：

*Hind*Ⅲ	*Bam*HⅠ
5′--------A A G C T T--------3′	5′--------G G A T C C--------3′
3′--------T T C G A A--------5′	3′--------C C T A G G--------5′
*Hind*Ⅲ 的识别序列	*Bam*HⅠ 的识别序列

又如 *Hin*f Ⅰ 的识别位点为：

5′--------G A N T C--------3′

3′--------C T N A G--------5′

限制性内切酶不仅识别的核苷酸序列不同，而且切割 DNA 后产生的片段的末端也是不同的。限制酶切割 DNA 后可以形成无单链突出的末端，即在同一水平切断 DNA 双链，称为平端或钝端（blunt end），如 *Sma* Ⅰ；也可以形成有单链突出的所谓黏性末端（cohesive end），即两链的切口错开 2～4 个核苷酸。黏性末端中有 5′突出的黏性末端，如 *Bam*H Ⅰ 水解产生的末端；也有 3′突出的黏性末端如 *Apa* Ⅰ 水解产生的末端。大多数酶是错位切割双链 DNA，产生 5′-或 3′-黏性末端，如图 5-1 所示。

5′--------C C C↓G G G--------3′ → 5′ G G G--------3′

3′--------G G G↑C C C--------5′ → 3′ C C C--------5′

Sma Ⅰ 的识别序列及切割部位　　切割后产生的平端

5′--------G↓G A T C C--------3′ → 5′----G A T C C--------3′

3′--------C C T A G↑G--------5′ → G--------5′

*Bam*H Ⅰ 的识别序列及切割部位　　切割后产生的 5′突出的黏端

5′--------G G G C C↓C--------3′ → C--------3′

3′--------C↑C C G G G--------5′ → 3′----C C G G G--------5′

Apa Ⅰ 的识别序列及切割部位　　切割后产生的 3′突出的黏端

图 5-1　几种限制酶切割后产生的平末端和黏性末端

Ⅱ 型限制性核酸内切酶的识别位点（靶顺序）可大致决定它切割 DNA 后产生的 DNA 片段的长短。对于识别 4 个核苷酸的内切酶（如 *Alu* Ⅰ），平均 256（4^4）个核苷酸长度就出现 1 个靶顺序，因此切割片段的长度约为 256bp；而 6 个特异核苷酸识别位点出现的概率为 1/4096（$1/4^6$），即切割片段长度为 4096bp。

（二）同工异源酶与同尾酶

来源不同但能识别和切割同一位点的酶称为同工异源酶（isoschizomer），也称同裂酶。如 *Bam*H Ⅰ 和 *Bst* Ⅰ（G↓GATCC）是一对同工异源酶，*Xho* Ⅰ 和 *Pae*R7（C↓TCGAG）也是一对同工异源酶，这些同工异源酶可以相互替代。

有一些同工异源酶对于切割位点上的甲基化碱基的敏感度有所不同，如 *Hap* Ⅱ 和 *Msp* Ⅰ 是一对同工异源酶，它们共同的靶序列为 CCGG。当序列中出现一个甲基化的胞嘧啶时（CC* GG），*Hap* Ⅱ 就不能够切割它，而 *Msp* Ⅰ 则不受影响。这样细微的差异可以用来研究 DNA 的甲基化作用。

有些限制酶识别序列不同，但是产生相同的黏性末端，这些酶称为同尾酶（isoaudamers），如 *Bam*H Ⅰ（G↓GATCC）、*Sau*3A Ⅰ（N↓GATCN）、*Bcl* Ⅰ（T↓GATCA）、*Bgl* Ⅱ（A↓GATCT）等就是一组同尾酶，它们切割 DNA 后都形成由 GATC 4 个核苷酸组成的黏性末端，这些切割形成的末端可借碱基之间的互补作用相互连接，使 DNA 重组具有更大的灵活性。应当注意的是同尾酶切割 DNA 产生的末端连接后所形成的新的位点（通常称杂交位点，hybrid site），有些可以被原来的同尾酶所识别，有些则不能识别，如图 5-2 所示。

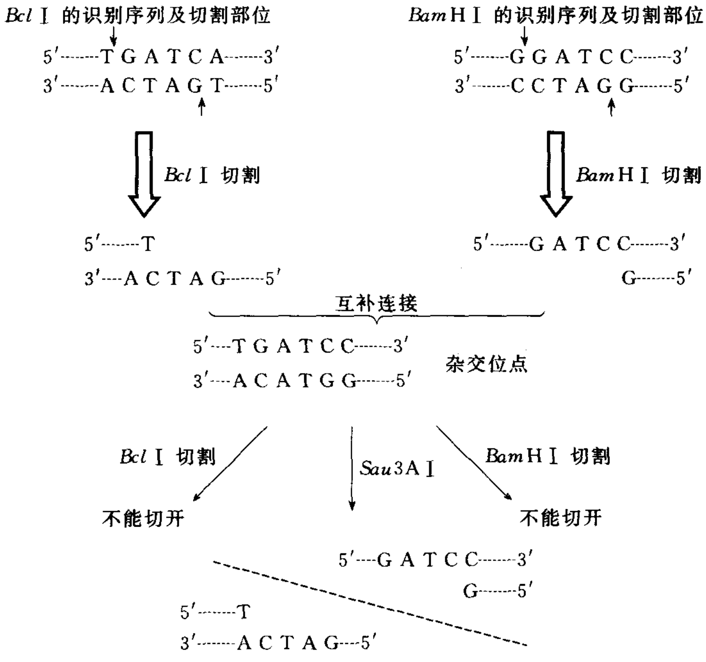


图 5-2 杂交位点的形成及其切割

（三）限制性核酸内切酶活性及其影响因素

限制性核酸内切酶活性大小体现在酶对 DNA 的水解程度不同，常用酶的活性单位来衡量其活性大小。酶的活性单位是指 1μg 纯 DNA 在指定缓冲液中，于 37℃孵育 1h 完全酶解 DNA 中所有同一限制性内切酶位点所需要的酶量。影响限制酶反应的因素很多，DNA 制品中的污染物（如蛋白质、酚、氯仿、乙醇、EDTA、SDS、高盐浓度）及缓冲条件等，均能抑制酶的活性。这种抑制可通过增加酶量（10～20U/μg DNA）、增大反应体积及适当延长反应时间来加以克服。限制性内切酶需要在－20℃低温下保存。商品化的限制性内切酶大多使用 50％的甘油缓冲液作为酶的储存液以避免酶蛋白的反复冻融。在酶切反应中，应该将酶的体积限制在反应总体积的 1/10 以内，否则酶活性会因为甘油浓度太大而受到抑制。

二、DNA 连接酶

常用的 DNA 连接酶（DNA ligase）有 T4 DNA 连接酶和大肠杆菌 DNA 连接酶 2 种。DNA 连接酶可使一段 DNA 的 3′-羟基末端和 5′-磷酸末端形成 3′,5′-磷酸二酯键，把两个 DNA 片段连接在一起，封闭 DNA 双链上形成的切口。连接酶同限制酶一样是基因克隆技术中不可缺少的重要工具酶。两种连接酶中，最常使用的是 T4 DNA 连接酶，它的连接效率高，既可用于黏性末端的连接，也可用于平末端的连接。虽然 T4 连接酶的最佳反应温度是 37℃，但由于在这个温度下，黏性末端之间氢键的结合不够稳定，故通常在低温下进行连接反应。一般说来，片段越小，末端黏性越强，连接反应则可使用较高的温度。DNA 平末端的连接效率比黏性末端的连接效率低得多。平端连接一般在 10～20℃进行，且需要较高的 DNA 浓度和 T4 连接酶的浓度（约比黏端连接高 10～20 倍）。

连接酶催化的连接反应是一种耗能的反应过程。在大肠杆菌 DNA 连接酶催化的连接反应中所需的能量由 NAD^+ 提供，而在 T4 DNA 连接酶催化的反应中，则由 ATP 提供。除此

以外，两种酶在其他方面的作用机制相同。

应当注意的是，DNA 连接酶并不能催化两条单链 DNA 分子的连接，而只能连接双链 DNA 分子的单链切口（nick），即在双链 DNA 中某一条单链上两个相邻核苷酸之间磷酸二酯键的断裂缺口。如果是双链中一个或多个核苷酸缺失所致的缺口（gap），DNA 连接酶不能连接。

三、DNA 聚合酶

DNA 聚合酶（DNA polymerase）催化以 DNA 为模板合成 DNA 的反应，在 DNA 重组技术中用于 DNA 的体外合成。经常使用的 DNA 聚合酶有大肠杆菌 DNA 聚合酶Ⅰ（全酶）、大肠杆菌 DNA 聚合酶Ⅰ的 Klenow 片段（Klenow 酶）、T4 DNA 聚合酶、耐高温 DNA 聚合酶，末端核苷酸转移酶以及逆转录酶等。这些 DNA 聚合酶的共同特点在于：它们都能够把脱氧核糖核苷酸连续地加到双链 DNA 分子引物链的 3′-羟基末端上，催化核苷酸的聚合，形成新的 DNA 链。

（一）DNA 聚合酶Ⅰ

在大肠杆菌中，DNA 聚合酶Ⅰ和聚合酶Ⅱ被认为主要参与 DNA 的修复，而 DNA 聚合酶Ⅲ则与 DNA 复制有关。在分子克隆操作中主要使用的是 DNA 聚合酶Ⅰ。

DNA 聚合酶Ⅰ是由大肠杆菌 *pol* A 基因编码的一种单链多肽，分子量 109000。它具有三种活性，即 5′→3′聚合活性、5′→3′外切活性和 3′→5′外切活性。

DNA 聚合酶Ⅰ的 5′→3′方向的外切活性，可以从 DNA 链的 5′-末端水解 DNA 分子。这种水解作用的主要产物是 5′-磷酸核苷，但同时也有少量的寡核苷酸片段，5′→3′外切酶活性可以随着 DNA 的合成而增强，而且对 DNA 双链中的单链缺口（带 5′-磷酸）也有活性。3′→5′方向的外切活性可以从 3′-末端开始向 5′-末端的方向水解 DNA，并释放出单核苷酸。酶水解反应的底物可以是单链也可以是双链 DNA 分子，但对双链 DNA，当有 dNTP 存在时，这种降解活性会被 5′→3′方向的聚合活性所抑制。

DNA 聚合酶Ⅰ的主要用途是通过 DNA 切口平移的方法制备 DNA 探针，用于核酸杂交分析。

（二）Klenow 酶

大肠杆菌 DNA 聚合酶Ⅰ全酶经枯草杆菌蛋白酶处理，可以产生大小不同的两个片段。其中的大片段称为 Klenow 片段（由 Klenow 等于 1970 年报道），分子量 76000。Klenow 聚合酶仍具有两种催化活性：①5′→3′聚合酶活性，可以合成 DNA；②3′→5′外切酶活性，可从 3′-端降解 DNA；但是它失去了 5′→3′外切酶活性。

在 DNA 分子克隆中，Klenow 酶的主要用途有：通过聚合作用填补或标记 DNA 分子的 3′-隐蔽末端，如图 5-3 所示；催化合成 cNDA 的第二链；DNA 序列测定。

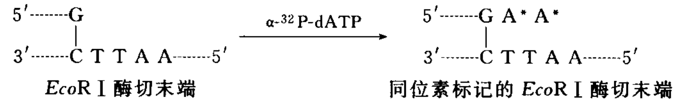


图 5-3 DNA 分子的 3 末端标记

（*表示同位素标记）

（三）*Taq* DNA 聚合酶

Taq DNA 聚合酶是第一个被发现的热稳定 DNA 聚合酶，分子量 65000，最初由 Saiki

等从温泉中分离的一株水生噬热杆菌（*Thermus aquaticus*）中提取获得。此酶能耐高温，在70℃反应2h后其残留活性大于原来的90%，在93℃反应2h后其残留活性是原来的60%，在95℃反应2h后其残留活性是原来的40%。在PCR过程中，由于*Taq* DNA聚合酶在变性步骤中（约94℃）不失活，可直接进入第二轮循环，因此不必每轮循环时重新加入新酶，这使得*Taq* DNA聚合酶成为PCR反应中独特用酶。

在*Taq* DNA聚合酶的基础上，目前已经发展多种新的耐热的聚合酶，如*Tfl* DNA聚合酶、*Tli* DNA聚合酶、*Pfu* DNA聚合酶等。其中较为突出的是*Pfu* DNA聚合酶，它具有3′→5′外切酶（校对）活性，是现今所描述过的所有热稳定酶中精确度最高的一种，其出错率大约是1×10^{-6}/bp，而*Taq* DNA聚合酶的出错率大约是1×10^{-5}/bp。但是，应当注意的是由于其天然的校对活性，所以扩增的PCR片段是平末端，适合于平端连接。或通过修饰，如在dATP存在下，使平末端片段与*Taq* DNA聚合酶共同保温，使其3′端带上一个A，以便与T-载体连接。另外，使用*Pfu* DNA聚合酶应注意在引物设计和PCR反应条件等方面与使用普通*Taq* DNA聚合酶的区别。一般建议在使用*Pfu* DNA聚合酶时，设计的引物应适当长些，PCR反应时采用热启动方式，其目的都是为了减少酶的3′→5′外切活性对引物的降解作用。由于*Pfu* DNA聚合酶比*Taq* DNA聚合酶的延伸反应速率低，因此PCR扩增时的延伸反应时间较长，约为每kb需2min，而*Taq* DNA聚合酶仅需要1min。

（四）逆转录酶

逆转录酶（reverse transcriptase）又称为RNA指导的DNA聚合酶。逆转录酶具有5′→3′聚合酶活性和RNase H活性。聚合作用所需模板可以是RNA，也可以是DNA。引物是带3′-羟基的RNA或DNA。RNase H活性是可以特异性地降解RNA-DNA杂交链中的RNA链。目前，已经从许多种RNA肿瘤病毒中分离到这种酶，但最普遍使用的则是来源于禽类成髓细胞瘤病毒（avian myeloblastosis virus，AMV）的逆转录酶和Moloney鼠白血病病毒（M-MLV）的逆转录酶。在体外以mRNA为模板合成cDNA是逆转录酶的最主要用途，此外，还可以用来对有5′突出末端的DNA片段作末端标记（补平反应）、用于双脱氧法测序和以DNA或RNA为模板合成核酸探针。

（五）T4 DNA聚合酶

T4 DNA聚合酶是从T4噬菌体感染的大肠杆菌培养物中纯化出来的一种特殊的DNA聚合酶。它是由噬菌体基因43编码，分子量114000，同大肠杆菌DNA聚合酶的Klenow片段相似，具有两种酶催活性，即5′→3′的聚合酶活性和3′→5′外切活性，而且3′→5′外切活性对单链DNA的作用比对双链DNA更强。

T4 DNA聚合酶可以用来标记DNA平末端或隐蔽的3′-末端。但反应必须在高浓度的dNTP存在下进行，以使5′→3′聚合反应压倒强劲的3′→5′外切活性，出现DNA净合成反应，从而补平缺口。对于3′-突出末端的DNA分子进行末端标记，可分两步进行。首先利用3′→5′外切活性切除3′-端突出尾，从而产生3′-凹端，然后在高浓度的某一种放射性同位素标记的dNTP存在下，外切降解反应与dNTP掺入的合成反应达到平衡，进而在T4 DNA聚合酶的聚合活性催化下，利用高浓度的dNTP逐渐取代的原有的核苷酸，这种反应称为取代合成。应用取代合成法可以给平末端的DNA片段或具有3′-隐蔽末端的DNA片段作末端标记。

与切口平移制备DNA探针的方法相比，T4 DNA聚合酶催化的取代合成法制备探针具有两个明显的优点：第一，不会出现切口转移法中可能产生的发夹结构；第二，经适当的限

制性核酸内切酶切割后可以很容易地转变成链特异性探针。

（六）T7 DNA 聚合酶

T7 DNA 聚合酶是从感染了 T7 噬菌体的大肠杆菌寄主细胞中纯化出来的一种复合形式的核酸酶，由两种不同的亚基组成：一种是 T7 噬菌体基因 5 编码的蛋白质，其分子量为 84000；另一种是大肠杆菌编码的硫氧还蛋白（thioredoxin），其分子量为 12000。T7 DNA 聚合酶具有极高的连续合成能力（即高度续进性），可以连续合成数千个核苷酸，是目前已知持续合成能力最强的 DNA 聚合酶，这种高度续进性得益于硫氧还蛋白的辅助作用，使得基因 5 蛋白与模板的亲合能力增强。除聚合活性以外，T7 DNA 聚合酶还具有单链及双链的 3′→5′核酸外切酶活性（由 T7 噬菌体基因 5 所编码）。类似于 T4 DNA 聚合酶，T7 DNA 聚合酶的 3′→5′核酸外切酶活性也很强，约为 Klenow 酶的 1000 倍。T7 DNA 聚合酶没有5′→3′外切酶活性。

T7 DNA 聚合酶有如下几个方面的用途：①用于长模板 DNA 的引物延伸反应；②通过单纯的延伸或取代合成的途径，标记 DNA 的 3′-末端。

T7 DNA 聚合酶经过适当化学处理可以使其失去 3′→5′外切酶活性而转变为一种新的 DNA 聚合酶——修饰的 T7 DNA 聚合酶。修饰的 T7 DNA 聚合酶由于失去了核酸外切酶活性，使其在单链模板上的聚合作用的速率增加了 3 倍，在物理特性、聚合酶性质以及连续合成能力等方面，都同天然的 T7 DNA 聚合酶完全一样。修饰的 T7 DNA 聚合酶是双脱氧终止法对长片段 DNA 进行序列测定的理想用酶，并以测序酶作为商品名广泛使用。此外，修饰的 T7 DNA 聚合酶也可用于有效地填补和标记具有 5′突出末端的 DNA 片段的 3′-末端。

四、修饰酶

在分子克隆操作中除了上述三类主要工具酶以外，还有一些其他的酶，可对 DNA 或 RNA 进行必要的修饰，以利于克隆的进行。如末端脱氧核苷酸转移酶可以催化一种或多种核苷酸连接到 DNA 片段的 3′-羟基末端形成同聚尾，可以用于寡核苷酸探针的末端标记；多聚核苷酸激酶可以催化人工合成的多核苷酸 5′-羟基末端的磷酸化，为连接反应提供磷酸基团或用于寡核苷酸探针的末端标记；碱性磷酸酶能特异地切除 DNA 或 RNA 的 5′-末端的磷酸基团，防止载体 DNA 分子自身环化；核糖核酸酶和脱氧核糖核酸酶在分子克隆技术中用于消化去除 RNA 或者 DNA 污染等。

（一）碱性磷酸酶

碱性磷酸酶有两种，分别来源于大肠杆菌和小牛肠道，前者叫做细菌碱性磷酸酶（bacterial alkaline phosphatase，BAP），后者叫做小牛肠碱性磷酸酶（calf intestinal alkaline phosphatase，CIP）。它们的共同特性是能够催化核酸分子脱掉 5′-磷酸基团，从而使 DNA（或 RNA）片段的 5′-P 末端转换成 5′-OH 末端（这一作用适合于 5′-突出末端、平端及凹端），这就是所谓的核酸分子的脱磷酸作用。

碱性磷酸酶的这种功能，对于 DNA 分子克隆是很有用的。例如，在 Maxam-Gilbert 序列分析法中，需要 5′-末端标记的 DNA 片段，为此必须在标记之前先从 DNA 分子上除去 5′-P 基团。另外，在 DNA 体外重组实验中，为了防止线性化的载体分子发生自我连接作用，也可以使用碱性磷酸酶，从载体片段上除去 5′-P 基团。

在实际应用中，主要使用小牛肠碱性磷酸酶（CIP）。这是因为 CIP 的比活性要比 BAP 高出 10～20 倍，同时 CIP 在 SDS 中加热到 68℃就可以完全失活，而 BAP 却是热抗性的酶，

所以要终止它的作用就很困难。为了去除极微量的 BAP 活性，需要用酚-氯仿反复抽提多次，远不如用加热法就可以使 CIP 完全失活来得方便实用。

（二）末端转移酶

末端脱氧核苷酸转移酶（terminal deoxynucleotidyl transferase），简称末端转移酶（terminal transferase），分子量为 60000，来源于前淋巴细胞和分化早期的类淋巴样细胞。在二价阳离子存在下，该酶能够催化 5′-脱氧核苷三磷酸进行 5′→3′方向的聚合作用，逐个地将脱氧核苷酸分子加到线性 DNA 分子的 3′-OH 末端。与 DNA 聚合酶不同，在反应中末端转移酶不需要模板的存在就可以催化 DNA 分子发生聚合作用。而且，它还是一种非特异性的酶，4 种 dNTP 中的任何一种都可以作为它的前体物，但对不同的核苷酸所需的阳离子略有不同。若加入的核苷酸为嘧啶核苷酸，则 Co^{2+} 为首选阳离子；若加入的核苷酸为嘌呤核苷酸，则 Mg^{2+} 为首选阳离子。另外，当反应混合物中只有一种 dNTP 时，就可以形成仅由一种核苷酸组成的 3′-尾巴，即同聚尾（homopolymeric tail）。

末端脱氧核苷酸转移酶催化作用的底物，即接受核苷酸聚合的受体 DNA，可以是具有 3′-OH 末端的单链 DNA，也可以是具有 3′-OH 突出末端的双链 DNA。然而，在含有 Co^{2+}、Mg^{2+} 或 Mn^{2+} 的低离子强度缓冲液中，这种酶同样也能够催化寡聚脱氧核苷酸的 3′-平端或 3′-凹端聚合上有限数量的核苷酸。

末端脱氧核苷酸转移酶的主要用途：①分别给外源 DNA 片段及载体分子加上互补的同聚物尾巴，以使它们可以连接起来；②用于同位素或非同位素标记 DNA 片段的 3′-末端，如将生物素-11-dUTP 掺入 3′-末端，便可作为荧光染料标记的抗生物素蛋白的接受位点。

（三）T4 多核苷酸激酶

T4 多核苷酸激酶（T4 polynucleotide kinase）来源于 T4 噬菌体感染的大肠杆菌细胞。它能催化 γ-磷酸从 ATP 分子转移给 DNA 或 RNA 分子的 5′-OH 末端。这种作用不受底物分子链的长短限制，甚至单核苷酸也同样适用。将 ATP 的 γ-磷酸直接转移到去磷酸化的 DNA 的 5′-OH 末端，这种反应被称为正向反应。另外一种磷酸转移的方式叫交换反应，在交换反应中，过量的 ATP 可使 T4 多核苷酸激酶将磷酸化 DNA 的 5′端磷酸转移给 ADP，然后 DNA 从 γ-^{32}P-ATP 中获得放射性同位素标记的 γ-磷酸而重新磷酸化。正向反应一般比交换反应常用、有效。

T4 多核苷酸激酶在 DNA 分子克隆中不仅可标记 DNA 的 5′-末端，而且还可以使缺失 5′-P 末端的 DNA 发生磷酸化作用。

（四）S1 核酸酶

S1 核酸酶来源于稻谷曲霉（*Aspergillus oryzae*），是一种高度单链特异的核酸内切酶，在最适的酶催反应条件下，降解单链 DNA 或 RNA，产生带 5′-磷酸的单核苷酸或寡核苷酸。对双链 DNA、双链 RNA 和 DNA-RNA 杂交体相对不敏感。通常水解单链 DNA 的速率要比水解双链 DNA 快 75 000 倍。这种酶需要低水平的 Zn^{2+} 的激活，最适 pH 范围为 4.0～4.3。一些螯合剂（如 EDTA 和柠檬酸等）能强烈地抑制 S1 核酸酶活性，此外磷酸缓冲液和 0.6%左右的 SDS 溶液也可以抑制它的活性，但它对尿素以及甲酰胺等试剂则是稳定的。

S1 核酸酶的单链水解功能可以作用于双链核酸分子的单链区，并从单链部位切断核酸分子，而且这种单链区可以小到只有一个碱基对的程度。由于具备这种功能，使 S1 核酸酶在分析核酸杂交分子（RNA-DNA）的结构、给 RNA 分子定位、测定真核基因中间隔子序列的位置、去除 DNA 片段中突出的单链尾，以及打开在双链 cDNA 合成期间形成的发夹环

结构等实验操作中，都是十分有用的工具。

第三节 基因克隆载体

重组的外源DNA必须进入宿主细胞中才能进行复制和表达。携带外源DNA进入宿主细胞，并在宿主细胞中进行无性繁殖或表达的小分子DNA称为载体（vector）。载体实际上就是DNA，按功能可分为克隆载体和表达载体两大类。克隆载体用于在宿主细胞中克隆和扩增外源DNA片段；表达载体则用于在宿主细胞中获得外源基因的表达产物（转录和翻译的产物）。

载体种类繁多，按其基本组成元件的来源不同，可以将载体分为质粒载体、噬菌体载体、噬菌粒载体、黏粒载体、病毒载体和人工染色体载体等。各类载体在大小、结构、性能等方面的特性差别很大，但作为DNA载体它们都具备以下一些共同的基本特点，这些特点构成了DNA载体所必须具备的基本条件。

① 具备自主复制能力，以保证重组DNA可以在宿主细胞内得到扩增；

② 具备较多的拷贝数，易与宿主细胞的染色体DNA分开，便于分离提纯；

③ 具备较小的分子量，易于操作，并有足够的接纳目的基因的容量；

④ 在非必需的DNA区段有较多的单一限制酶酶切位点，便于目的基因的克隆；

⑤ 有一个或多个筛选标记（如对抗生素的抗性、营养缺陷型或显色表型反应等）；

⑥ 具有较高的遗传稳定性。

目前可以满足上述要求的载体均为人工所构建，并有种类繁多的商品化的载体可供选择。每类载体都有它独特的结构与性质，适于不同的研究目的。本节简要介绍质粒载体、噬菌体载体及黏粒载体。病毒载体的有关内容参见第六章。

一、质粒载体

质粒是存在于细菌染色体之外的、具有自主复制能力的环状双链DNA分子。质粒的结构比病毒简单，它既没有蛋白质外壳，也没有细胞外生命周期，但是它能够在宿主细胞内独立增殖，并随着宿主细胞的分裂被遗传下去。在自然界中，质粒分布广泛，无论是真核还是原核生物细胞，也不论是革兰阳性菌还是革兰阴性菌，都已经发现质粒分子的存在。质粒作为一种广泛存在的、裸露的、能自主复制的简单DNA分子，非常适合于在基因克隆中作为外源DNA分子的载体在相应的宿主细胞内复制、传递、表达遗传信息，成为基因克隆操作中不可缺少的重要工具。

1. pBR322质粒载体

pBR322是一种经典的环状双链DNA质粒载体的代表，长4.36kb。在细菌中的分子个数（称为拷贝数，copy number）高，便于制备；含有一个复制起始点（*ori*），它可以保证这个质粒只在大肠杆菌中行使复制功能。含有*Bam*HⅠ、*Eco*RⅠ、*Hind*Ⅲ和*Sal*Ⅰ等多种限制性内切酶的单一酶切位点，用于插入外源DNA片段。

pBR322中有四环素（*tet*）和氨苄青霉素（*amp*）两个耐药性基因标记，分别简写为

tet^r 和 amp^r（r 是 resistance 的简写，即抗性或耐性）。无耐药性基因的大肠杆菌被 pBR322 质粒转化后便可获得相应抗生素抗性。两个抗生素抗性基因中均含有单克隆位点，如 *Bam*HⅠ、*Hind*Ⅲ和 *Sal*Ⅰ的识别位点在四环素耐药性基因内；而另一个单一的 *Pst*Ⅰ识别位点在氨苄青霉素耐药性基因内。若在这些位点插入外源 DNA 片段，破坏了相应的 tet^r 和 amp^r 基因，宿主细胞失去相应的耐药性，不能在含相应抗生素的培养基中生长，这一现象称为插入失活（insertional inactivation）。利用这一现象可用于筛选含重组 DNA 的宿主细胞。例如，将外源 DNA 片段克隆在 pBR322 质粒的 *Bam*HⅠ位点上，由于阻断了 tet^r 基因编码序列的连续性，而使其失去活性，结果便产生出了具有 amp^r tet^s（s 是 sensitivity 的简写，即敏感）表型的重组 pBR322 质粒。将这样的重组质粒转化野生型（amp^s tet^s）的大肠杆菌细胞，并涂布在含有氨苄青霉素的选择性培养基的平板上，那么存活下来的菌落都将具有 amp^r 表型。但在这些 amp^r 表型的转化子克隆中，有一部分具有 tet^r 表型，另一部分具有 tet^s 表型。因此，需要将它们再涂布在含有四环素的选择性培养基平板上，就可以迅速地辨别出它们到底属于什么表型。所以，重组体转化子克隆的筛选需经过两个步骤，即需从第一种抗性平板转移到第二种抗性平板。质粒 pBR322 的抗性基因筛选如图 5-4 所示。

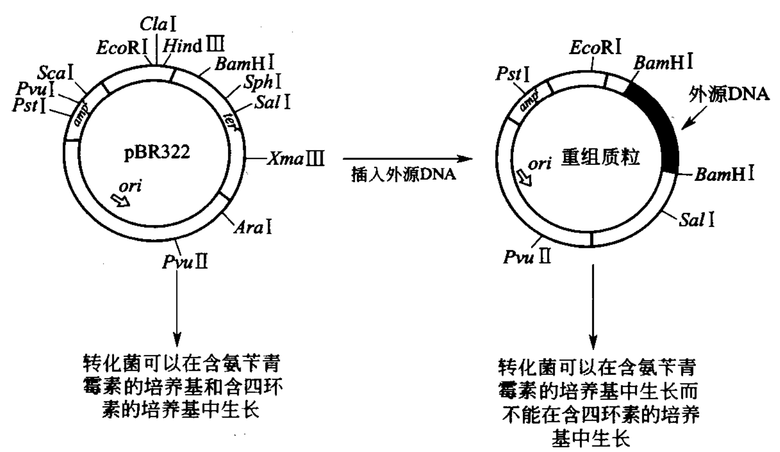


图 5-4　质粒 pBR322 的抗性基因筛选示意图

2. pUC18/19 质粒载体

pUC18/19 质粒载体是由大肠杆菌 pBR 质粒与 M13 噬菌体改建而成的环状双链 DNA 克隆载体。pUC18/19 长 2.69kb，复制起始点来自 pBR322，含有一个氨苄青霉素抗性基因、一个大肠杆菌 β-半乳糖苷酶基因（*lac*Z 基因）启动子及一个调节 *lac*Z 基因表达的阻遏蛋白（repressor）的基因 *lac*Ⅰ，还含有 MCS 区段。pUC18 和 pUC19 所有元件均相同，两者唯一的差别是相对于 *lac*Z 启动子的多克隆位点的排列方向相反，提供了更多的克隆策略选择机会。

pUC18/19 质粒载体含有 *lac*Z 基因，能表达 β-半乳糖苷酶氨基端的一个片段（α 片段），适用于可编码 β-半乳糖苷酶羧基端的部分序列（片段）的宿主细胞（DH5α，JM103，JM109 等）。虽然这类质粒和宿主各自所编码的片段都没有完整的 β-半乳糖苷酶的活性，但两者可以互补，产生具有完整 β-半乳糖苷酶活性的蛋白质，这种现象称为 α 互补（alpha complementation）。将这些细胞涂布在含有氨苄青霉素、IPTG 和 X-gal 的固定培养基上，那些带有自身环化质粒的细胞由于能够形成有活性的 β-半乳糖苷酶，所以菌落是蓝色的，如

果 pUC18/19 质粒中插入有目的基因片段，那么就会破坏 *lac*Z 基因的结构，导致细胞无法产生功能性的 *lac*Z 蛋白，也就无法形成杂合 β-半乳糖苷酶，因而菌落是白色的。利用这一特性建立了一个简单的颜色实验用于重组体的筛选。

3. pGEM 系列载体

pGEM 质粒载体的典型代表为 pGEM-3Z 和 pGEM-4Z，它们与 pUC 质粒载体十分相似。但 pGEM-3Z/4Z 具有两个来自噬菌体的启动子，即 T7 和 SP6 启动子，具有 RNA 聚合酶的特异性识别位点。当在 T7 或 SP6 RNA 聚合酶存在条件下，可以在体外转录出克隆在载体中的外源基因的 mRNA。pGEM-3Z 和 pGEM-4Z 作用相似，主要差别在于 T7 和 SP6 启动子位置互换、方向相反，如图 5-5 所示。

4. 穿梭质粒载体

穿梭质粒载体（shuttle plasmid vector）是指一类由人工构建的、具有两种不同复制起始点和选择性标记，因而可在两种不同的宿主细胞中存活和复制的质粒载体。通常，穿梭质粒既能在原核细胞中复制，又能在真核细胞中复制和表达，即来往于两种不同的宿主，称为穿梭。例如大肠杆菌和酵母菌之间的穿梭质粒 YEp，它具有大肠杆菌中复制的起始序列和抗性基因，也具有酵母菌中复制的序列和营养标记基因。在酵母中使用的质粒一般都是穿梭型的，因为从酵母中提取质粒比较困难，通常将这类质粒放在大肠杆菌中保存、提取和扩增，构建成重组质粒后才转入酵母细胞内。

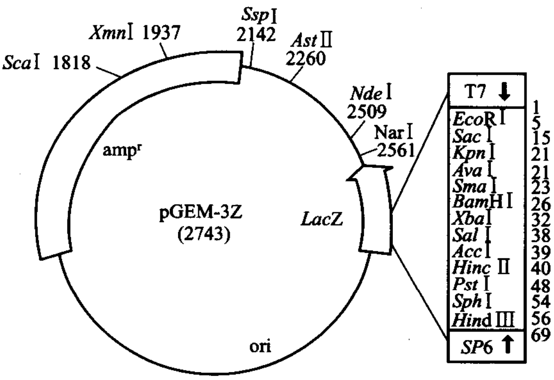


图 5-5　pGEM-3Z 质粒载体示意图

在动物体系中也已经发展了类似的穿梭质粒载体。现在，已经有大肠杆菌质粒与牛乳头瘤病毒、腺病毒、逆转录病毒的穿梭载体，如 pBPV-BV1 就是一种典型的动物细胞穿梭载体，它既可以在大肠杆菌中复制，也可以在动物细胞中复制和表达。有人用这种载体克隆人 β-干扰素基因，在大肠杆菌中操作，然后在动物细胞中表达，取得了满意的效果。

二、噬菌体载体

噬菌体（phage）是一类细菌病毒的总称，即感染细菌的病毒。它的结构与细菌或真核细胞相比，显然十分简单，但与质粒相比，则要复杂得多。不同种类的噬菌体颗粒，在结构上有很大的差别，主要有双链噬菌体和单链丝状噬菌体两大类。前者为 λ 噬菌体类，后者包括 M13 噬菌体和 f1 噬菌体。常用作基因工程载体的噬菌体有 λ 噬菌体和 M13 噬菌体。

1. λ 噬菌体 DNA 结构

λ 噬菌体的基因组是一长约 48kb 的双链 DNA，在宿主体外与蛋白质结合，包装成为含有双链线状 DNA 分子的颗粒。λ 噬菌体的基因组包括 50 个基因，调控非常复杂。除了两个正调节基因外，其余的基因都是按功能的相近性聚集成簇的，例如头部、尾部、复制及重组等功能区，如图 5-6 所示。

在 λ 噬菌体线性双链 DNA 分子的两端，各有一条由 12 个核苷酸组成的彼此完全互补

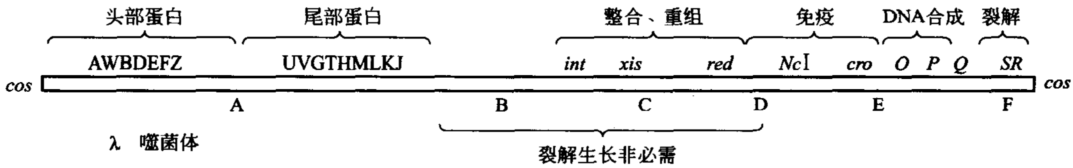


图 5-6 λ 噬菌体的基因分区

的 5′单链突出序列，即通常所说的黏性末端。注入到寄主细胞内的 λ 噬菌体的线性 DNA 分子，会迅速地通过黏性末端之间的互补作用，形成环形双链 DNA 分子。随后，在 DNA 连接酶的作用下，将相邻的 5′-P 和 3′-OH 基团连接起来。这种由黏性末端结合形成的双链区段称为 cos 位点（cohesive-end site），即黏性末端位点，如图 5-7 所示。

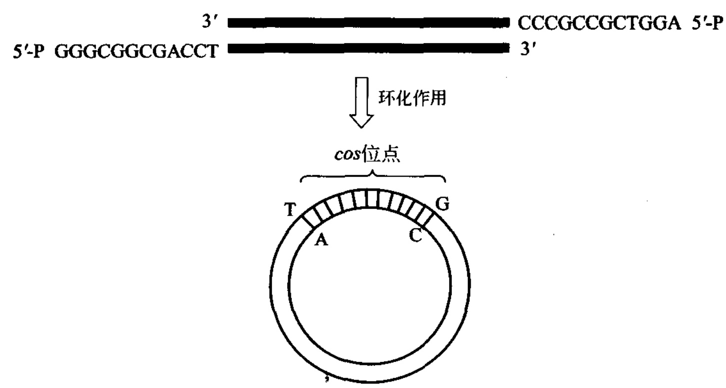


图 5-7 λ 噬菌体线形 DNA 分子的黏性末端及其环化

2. λ 噬菌体载体的构建

野生型 λ 噬菌体经过改造，已衍生出 100 多种克隆载体，目前常用的有 EMBL 系列、λgt 系列和 Charon 系列。λ 基因组中只有 60%（约 30kb）是溶菌生长所必需的，中间 40%的区域是非必需的，用作克隆载体时，可以将这部分切除，插入外源 DNA 片段。这种载体称为取代（置换）型载体（replacement vector）。取代型载体可以允许插入的外源 DNA 大小为9～23kb。如无外源 DNA 取代，由 λDNA 左右臂连接起来的 DNA 分子，由于太小而不能包装，这样就提供了一个挑选重组 λDNA 的阳性标志。

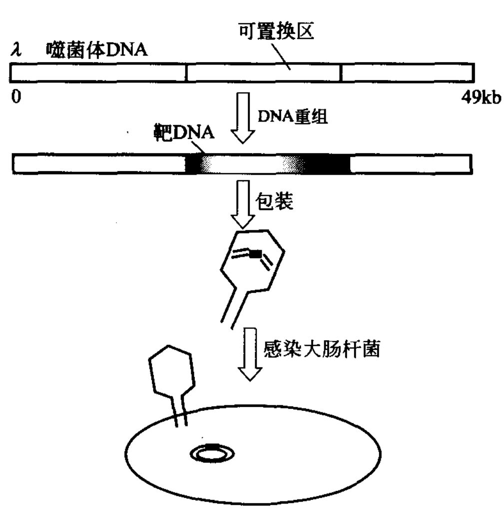


图 5-8 置换型 λ 噬菌体克隆示意图

取代型载体最常见的是 EMBL 系列和 Charon 系列，这些载体在 λ 噬菌体中央可替换区两侧各有一个或多个酶切位点，用相应的酶切割后，外源 DNA 片段可置换 λ 噬菌体非必需区序列而插入载体，如图 5-8 所示。由于它允许插入的外源 DNA 片段长度大，因而适用于基因组 DNA 的克隆及基因组 DNA 文库和 cDNA 文库的构建。λ 噬菌体感染细菌后，其 DNA 可在宿主菌内大量复制，并表达出包装子

代噬菌体所需要的蛋白质，进而包装出大量的子代噬菌体，最后发生溶菌作用，释放出病毒。

在λ噬菌体的整个生命周期中，处于溶原生长状态时形成混浊的噬菌斑，处于溶菌生长状态时则形成清晰的噬菌斑。一旦外源 DNA 插入某些区段，可使λ噬菌体中编码限制酶活性阻遏物的基因被分割而失去功能（插入失活），而不能进入溶原生长。因此，可根据噬菌斑的外观差别将体外重组的λ噬菌体同亲本λ噬菌体区别开来。

除置换型载体以外，还有一种插入型载体，如λgt（λgt10、λgt11 等）系列，λgt10 的外源 DNA 插入位点 *Eco*RⅠ处于阻遏物基因上，外源 DNA 插入后使阻遏物基因失活，这样形成的重组噬菌体所感染的大肠杆菌呈现透明的噬菌斑，而未重组者形成浑浊的噬菌斑，很容易区分。这类载体适用于 6～8kb 大小的 DNA 片段的插入。λgt11 载体含有 *lacZ* 基因，插入位点 *Eco*RⅠ在 *lacZ* 基因的 3′-端，外源基因插入后（即重组的λgt11）导致 *lacZ* 基因失活，在含有 X-gal 培养基中形成白色噬菌斑，而未重组的λgt11 则形成蓝色斑点，易于区分。另外，当插入的 cDNA 的阅读框与 *lacZ* 基因相一致时，能产生融合蛋白质，可以用免疫学方法，如蛋白质印迹分析进行检测。这类载体常用于 cDNA 的克隆或 cDNA 文库的构建。

三、黏粒载体

λ噬菌体的实际克隆容量不超过 25kb，但一些真核生物的基因很大，如鸡胶原蛋白基因长度达 38kb，一般载体不能克隆这么大的基因，于是设计构建了一种能克隆真核生物基因的载体，称为黏粒（cosmid），是黏性质粒的简称，又称为柯斯质粒。它是一类人工构建的含有质粒复制子和λ噬菌体的 *cos* 黏性末端序列的特殊类型的质粒载体。*cos* 位点的存在可以使 DNA 包装入噬菌体的头部。黏粒中含有质粒的复制起始位点、一个或多个限制性内切酶位点、耐药性基因标记和λ噬菌体的 *cos* 黏性末端序列，兼有λ噬菌体和质粒两方面的优点，大小约为 4～6kb，允许克隆的外源 DNA 片段长度为 31～45kb，而且能被包装成为具有感染能力的噬菌体颗粒。常用的黏粒有 pJ 系列和 pH 系列，如 pHC79 黏粒就是由λ噬菌体片段与 pBR322 质粒构建而成。黏粒主要用于真核细胞基因组文库的构建。

第四节　基因克隆的一般方法

克隆基因的基本步骤包括：获得含目的基因的 DNA 片段；外源 DNA 片段与载体的体外连接，构建重组体分子；重组体导入宿主细胞，并在宿主细胞内复制和扩增；重组体的筛选与鉴定。

上述过程中的首要任务就是要获得目的基因片段。获得目的 DNA 片段的方法很多，主要有：人工化学合成、逆转录制备 cDNA、构建基因组文库（genomic library）或 cDNA 文库（cDNA library）、PCR 方法以及各种差异显示方法等。

一、获得目的基因的常用方法

（一）化学合成法

化学合成法是指通过化学的方法人工合成所需要研究的 DNA 序列的方法。目前，随着

合成技术的发展，应用计算机控制的全自动核酸合成仪可以按设计好的序列一次合成100～120bp长的DNA片段。因此一定长度的基因片段是可以利用化学方法合成的，但前提条件是已知目的基因的核苷酸序列，或能根据基因产物的氨基酸顺序推导出编码基因的核苷酸序列。对于分子较大的基因，可以先分段合成DNA短片段，然后按一定顺序连接短片段，组装成一个完整的基因。目前常用的组装方法主要有以下两种。

一是短片段直接连接法。这种方法是先将化学合成的一条寡核苷酸片段激活，使其带上必要的5′-磷酸基因，然后再与另一条部分互补的寡核苷酸片段退火，形成带有黏性末端的双链寡核苷酸片段。把这些双链寡核苷酸片段混合在一个试管中，加上T4 DNA连接酶，使它们彼此连接，组装成一个完整的基因或基因的一个片段。应用这种基因组装法已经构建出了许多基因，其中包括分子量较大的α-干扰素和牛视紫红质基因。

另一种方法是合成较长片段经部分重叠后连接。这种方法可以减少合成的寡核苷酸片段的数量，它通过合成两条具有互补3′-末端的长的寡核苷酸片段，混合后退火，两片段3′-端互补区配对重叠，加入DNA聚合酶（如Klenow酶）后，在聚合酶作用下，迅速地合成出相应的互补链。理论上讲，应用这种方法能够合成出大分子量的基因，并可以任意制造、修饰基因。然而，实际应用上却存在困难，这是因为：①退火过程中难以消除那些不必要的杂交作用，从而产生非特异性产物；②大分子量的基因结构复杂，不可避免地存在二级结构，从而干扰聚合酶的作用；③突变频率相当高，平均每隔500～800bp就会出现一个突变。因此这种基因组装法还是有相当的局限性。

（二）逆转录PCR克隆目的基因

逆转录PCR（RT-PCR）方法获得目的基因是一种较为简单实用的方法，它可以不用构建cDNA文库，但需要知道该基因的cDNA或mRNA的序列信息，因此该方法适合于对已知基因的克隆或不同物种的高度保守基因的克隆。逆转录PCR克隆方法的出发原料是细胞的总RNA或mRNA，因此必须首先从细胞中分离总RNA或mRNA。总RNA的分离提取方法较多，提取相对容易；mRNA的分离提取相对困难一些，一般是在提取总RNA的基础上，再用oligo(dT)纤维亲和层析的方法提取。值得注意的是，细胞中mRNA的丰度对这种克隆方法具有重要影响。高丰度的mRNA克隆较为容易。对于低丰度mRNA，则在具体克隆操作中会遇到很大的困难。由于丰度低，mRNA在提取和逆转录过程中很容易损失或被降解。另外，在PCR扩增中，模板太少也不容易获得期望的目的基因片段。

可以采用三种方法合成相应的cDNA作为PCR反应的模板，这三种方法所用引物分别为：基因特异性引物（gene specific primer，GSP），多聚寡核苷酸引物［oligo(dT) primer］，随机六核苷酸引物（random hexamer primer）。

用随机引物逆转录方法的特异性最低。随机引物可以与mRNA上多个位点配对结合，产生短的、部分cDNA片段，因此该方法适合于获得mRNA 5′-端序列或具有二级结构的RNA模板序列，不适合全长cDNA序列的克隆。

用oligo(dT)作引物进行逆转录的方法特异性高于随机引物的方法。由于oligo(dT)可以和大多数真核生物mRNA 3′-端的poly(A)结合，因此该方法适合于大部分真核mRNA的逆转录反应。常用的oligo(dT)为长度为12～18个脱氧胸苷酸（T）组成的多聚T。

基因特异性引物是特异性最高的逆转录方法。所谓基因特异性引物（GSP）是指能够与特定RNA目标序列（靶序列）互补（杂交）的寡核苷酸片段。GSP的设计符合一般引物设计原则，通常设计在mRNA的3′-端附近，设计时应注意其序列应与mRNA互补。

（三）PCR 方法克隆目的基因

利用 PCR 方法克隆目的基因比传统的 DNA 克隆方法简单、方便，但它一般要求知道目的基因的碱基序列或至少知道其两侧的部分序列，这样才能设计合适的 3′-端或 5′-端引物序列，进行 PCR 扩增反应，目前常用的方法有套式 PCR（也称巢式 PCR，nested PCR）、反向 PCR（inverse PCR）、锚定 PCR（anchored PCR，又称单侧 PCR）等，详见 PCR 章节内容。

（四）利用基因表达差异的克隆方法

基因表达具有时空特异性，在一定的发育时期，某一特定细胞内，只有约 15%的基因得以表达。通常将这种按不同时间、空间进行有序的表达方式，叫做基因的差别表达。利用基因的表达差别，近 10 年来不断发展了一些新的克隆方法，使克隆基因的数量大为增加，克隆效率迅速提高。通常，可以一次发现或克隆多个差异表达的基因片段，结合基因定位技术和生物信息技术，可以迅速使差异表达的基因片段与其染色体上的完整基因联系起来，这已成为目前克隆新基因的重要方法。

1992 年 P. Liang（梁鹏）和 A. D. Pardee 发明的 mRNA 差异显示（mRNA differential display，DD）或称为差别显示逆转录 PCR（differential display reverse transcription PCR，DDRT-PCR），能有效地鉴定并分离出差别表达的基因。经过 10 多年的实践，DDRT-PCR 现已成功地分离到数百种基因，成为分离编码产物未知的目的基因的一种十分快速而有效的方法之一。

1. DDRT-PCR 的基本原理和方法

现已知道，真核基因 mRNA 分子的 3′-末端都带有多聚腺苷酸尾巴，即通常所说的 poly（A）尾。利用这一特定序列结构，以 oligo(dT) 为引物可以合成出与 mRNA 互补的 cDNA。仔细分析 mRNA 分子 3′-端序列结构发现，在 poly(A) 序列的起点碱基之前（亦即 5′上游）的碱基，除了碱基为 A 的情况之外，只能有如下 3 种可能的排列组合，如图 5-9 所示。根据这种序列结构特征，P. Liang 等人设计合成了总共 12 种不同的下游引物，用以逆转录 mRNA，合成 cDNA 第一链。这种引物通常叫做锚定引物（anchored primer），它是由 11 个或 12 个连续的 T 加上两个 3′-端的特定脱氧核苷酸（又称锚定脱氧核苷酸）组成，并用 5′-T_{11}MN 或 5′-T_{12}MN 通式表示。其中 M 是除了 T 以外的任何一种核苷酸（即 A、G

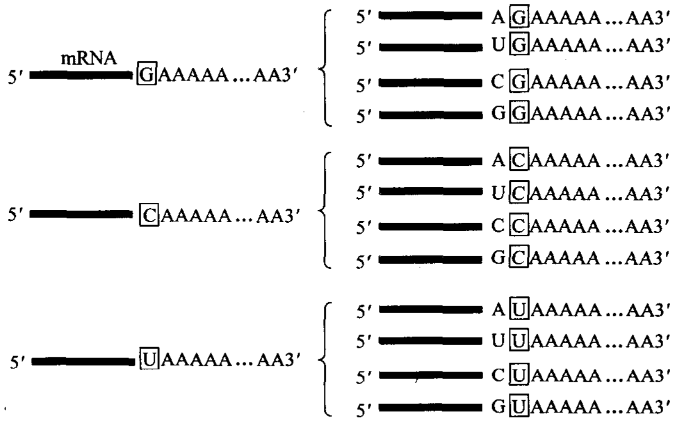


图 5-9 poly(A) 尾部序列结构

或 C)，而 N 则是任何一种核苷酸（即 A、G、C 或 T)，故 MN 共有 12 种不同的排列组合方式。这样，每一种此类锚定引物，都将能够把总 mRNA 群体的 1/12 逆转录成 mRNA-cDNA 杂交分子，即将整个 mRNA 群体在 cDNA 水平上，分成数目大致相等的但序列结构不同的 12 份亚群体。

为了展示这些 cDNA 片段，就需要通过 PCR 扩增。因此，在 DDRT-PCR 反应中，还需使用另外一种由 10 个核苷酸组成的 5′-端随机引物。这种 5′-端随机引物，可以随机地与 mRNA 的不同部位结合，当与锚定引物组成引物对，以第一链 cDNA 为模板进行 PCR 扩增时，在标准的序列胶电泳中，可以显示出 50～100 条长度在 100～500bp 之间的 DNA 条带。显然，如果使用 12 种 3′-端锚定引物和 20 种 5′-端随机引物组成的全部 240 组引物对分别做 PCR 扩增，就应能产生出 20000 条左右 DNA 条带，其中每一条都代表一种特定的 mRNA 种。这个数字涵盖了一定的发育阶段某种类型细胞中所表达的全部的 mRNA 种。

如果在 PCR 反应体系中加入同位素标记的 dNTP（通常用 ^{32}P 标记），PCR 扩增产物经序列胶电泳后 X 光胶片曝光，则可以在胶片上显示扩增条带。通过对比两种细胞的扩增结果，即可获得差异表达条带，如图 5-10 所示。

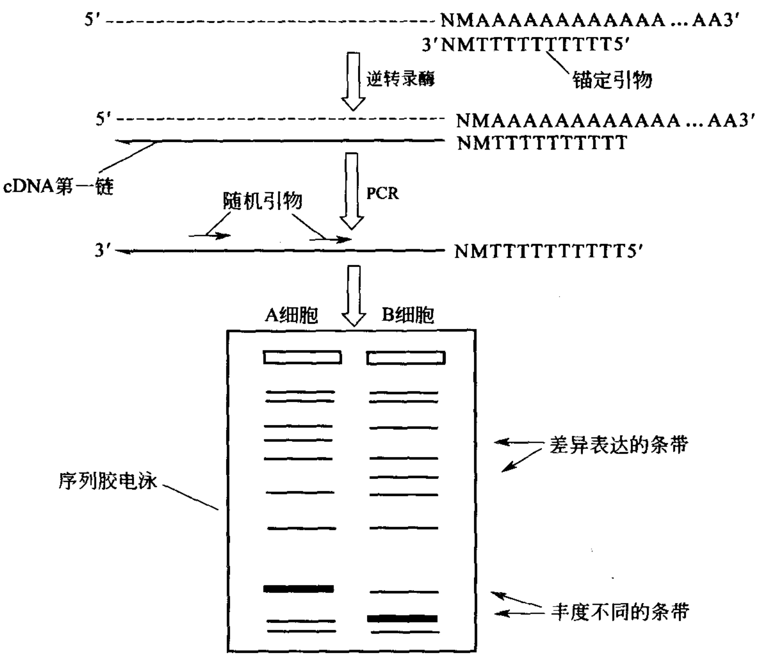


图 5-10　DDRT-PCR 反应基本程序

2. DDRT-PCR 的特点

与其他常用方法相比，DDRT-PCR 具有下列优点：①简单方便，它只用 PCR 和序列胶电泳两项技术；②灵敏度高，5μg 的总 RNA 可以满足所有引物组合 PCR 扩增的需要，一些低丰度 mRNA 也可以被检测出来；③重现性好，99％的条带可以重复显示；④可同时比较多种细胞类型，也同时展现所有差异；既可以检测基因表达的上调，也可以检测表达下调；⑤可以检测新基因。

经过多年的实践，人们也发现 DDRT-PCR 方法存在一些明显的缺点。其中突出的问题

是假阳性比例较高（可高达50%～75%）、PCR组合数目多工作量大。经过多年研究改进，目前已经将3′-端引物简化为$T_{11}G$、$T_{11}A$和$T_{11}C$，这大大简化了实验的步骤，同时不影响实验结果。另外，在实验装置上也有改进，提高了分辨率。对假阳性过高的问题也有改进，Callard等在回收cDNA时，先将回收片段的一半进行PCR，然后克隆到适当的载体中，另一半制成探针，通过斑点印迹筛选以上克隆，有效地去除了DDRT-PCR方法所产生的假阳性结果。在应用过程中还有新的发现和其他方面的改进，目的都是使该方法更简化实用。目前已有公司专门针对DDRT-PCR设计了第二代、第三代产品，使该方法逐渐完善，成为新基因克隆的主要方法之一。

二、目的基因与载体连接

目的基因分子同载体的连接主要依赖限制性核酸内切酶和连接酶的作用。连接方式很多，应根据克隆工作的具体情况灵活选用。一般要求在不影响外源基因密码结构的基础上尽可能选择简单易行的连接方式，并能回收外源基因片段。具体连接方式分两大类，即黏端连接和平端连接。

（一）具黏性末端的DNA片段之间的连接

1. 互补黏端之间的连接

当外源DNA与载体用同一种限制酶消化（酶切），则可以产生互补的黏性末端。将具有互补黏端结构的载体与目的基因混合后退火，在DNA连接酶作用下即可将两者连接起来，如图5-11所示。

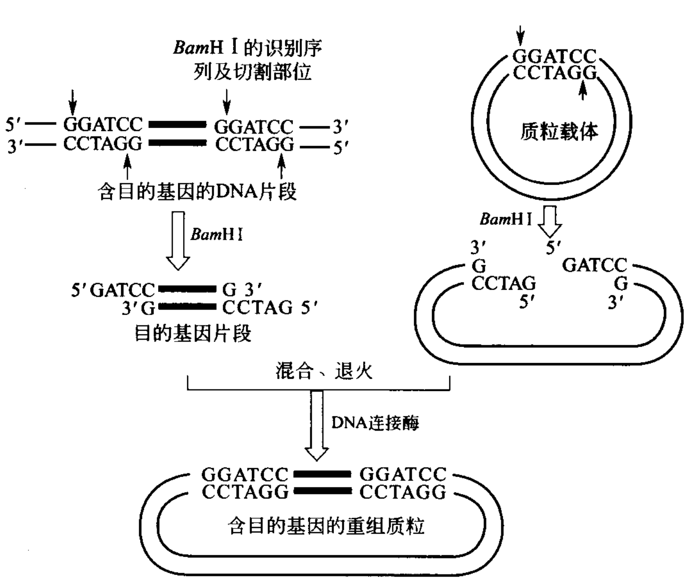


图5-11　同一种限制酶切割DNA产生的黏性末端的连接

在上述连接反应中，载体和目的基因是用同一种限制酶切割后连接的，是一种单酶切的连接方式。在连接后形成的重组分子中，目的基因可以有两种不同的插入方向，即正向插入和反向插入。两种插入方向对于基因克隆来说并无影响，但如果是为了表达外源基因则只能选用正向插入的重组体（即目的基因位于载体中启动子的下游）。

如果对载体和目的基因都分别采用两种不同的限制酶切割，如用*Eco*R Ⅰ和*Hind*Ⅲ分别

切割载体和目的基因，则目的基因酶切片段的两端可以分别与载体分子相应的酶切末端互补连接。由于两种酶的酶切末端不同，因此，酶切产生目的基因片段只能以一种特定的方向插入载体分子，如图 5-12 所示。

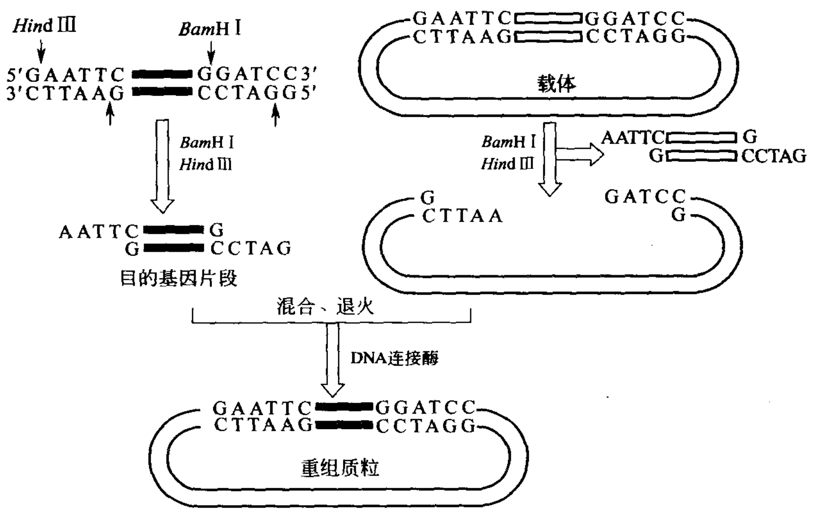


图 5-12　双酶切 DNA 片段黏性末端的连接

2. 非互补黏端之间的连接

目的基因片段与载体分子之间并不总是互补的黏性末端，经常会遇到非互补的黏端和平末端。平末端的连接方式很多，将在下面专门讨论。非互补的黏端是不能用 T4 DNA 连接酶连接的，常用的方法是将其转化为平末端，如用 S1 核酸酶水解突出末端或用 DNA 聚合酶补平 3′-凹端，再按平末端的连接方法连接。

（二） 平末端的连接

有些限制酶水解 DNA 产生平头末端，如 *Alu* Ⅰ、*Bal* Ⅰ 等。平末端 DNA 片段比较难连接，在高浓度的 T4 DNA 聚合酶催化下可以被有效连接，只是连接效率明显低于黏端之间的连接，并且连接之后不能从连接处切除下来。平末端片段除了直接连接以外，更常用的方法是将平末端转化为互补黏端，提高连接效率。目前常用的转化方法主要有两种，即同聚物加尾法和连接物连接法。

1. 同聚物加尾法

同聚物加尾法是平末端 DNA 片段连接的有效方法。对于待连接的 DNA 片段，利用末端转移酶的催化作用，在平末端的 3′-OH 上逐个聚合上一段寡聚脱氧核苷酸，例如寡聚 G，即 oligo(dG)；在另一片段（如线性化的载体）上则聚合上一段与上述末端互补的寡聚脱氧核苷酸，如寡聚 C，即 oligo(dC)。由于 G-C 间的互补关系，使得平末端转化为互补的黏性末端，从而可以很方便地连接，如图 5-13 所示。

实际上 DNA 末端聚合上去的 oligo(dG) 和 oligo(dC) 的长度不会完全一致，但这并不妨碍其互补作用，如果在连接体系中加有 DNA 聚合酶及 dNTP，则可在 DNA 聚合酶催化下填补互补留下的空缺，剩下的切口（nick）由连接酶连接封口。

2. 运用连接物

连接物连接法泛指通过连接物将平末端 DNA 片段连接起来的方法。连接物有两种：接

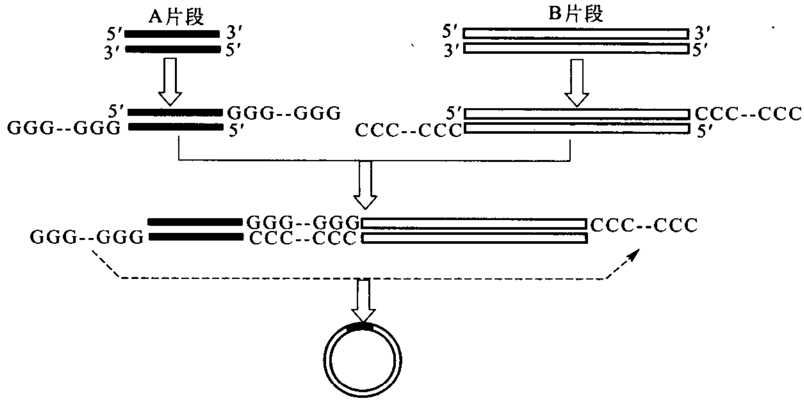


图 5-13 同聚物加尾连接示意图

头（linker）和衔接头（adaptor）。

接头是指化学合成的一段由 10～12 个脱氧核苷酸组成的具有一个或数个限制酶识别位点的平末端双链寡核苷酸短片段。用接头连接时，先将 DNA 片段的 5′-端和接头的 5′-端用多核苷酸激酶处理使之磷酸化，再通过 T4 DNA 连接酶将两者连接起来。随后，用相应限制酶消化接头和载体分子，产生互补的黏性末端，即可按上述互补黏端的连接方法连接，如图 5-14 所示。

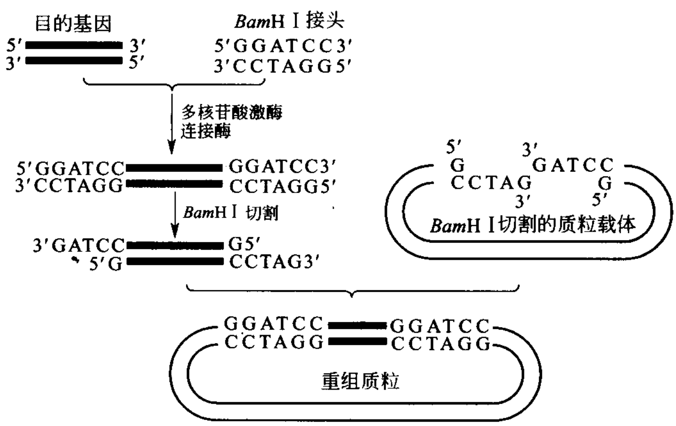


图 5-14 连接物分子连接平末端的 DNA 片段示意图

衔接头与接头略有不同，它是一头为平末端另一头具有某种限制酶黏性末端的双链寡核苷酸短片段。当衔接头与外源 DNA 分子连接后便形成了带有黏性末端的新的 DNA 分子，可与经同样限制酶切割的载体分子互补连接。

三、构建文库克隆目的基因

（一）构建 cDNA 文库克隆目的基因

以 mRNA 为模板，经逆转录酶催化合成 DNA，其序列与 mRNA 互补，故称为互补 DNA（complementary DNA，cDNA）。构建 cDNA 文库时，提取出组织细胞的全部 mR-

NA，在体外逆转录成cDNA，与适当的载体（常用噬菌体或质粒载体）连接后转化受体菌，则每个细菌含有一段cDNA，并能繁殖扩增，这样包含着细胞全部mRNA信息的cDNA克隆集合称为该组织细胞的cDNA文库。对于真核生物来说，编码蛋白质的基因约有5万个，但在特定发育阶段的组织细胞内只有约15%的基因得以表达，产生mRNA。由mRNA出发的cDNA克隆，就其复杂程度而言，比从基因组克隆简单。

cDNA文库的构建过程归纳为4个步骤，包括一系列酶促反应，分述如下。

1. 细胞中mRNA的分离纯化

构建cDNA的第一步是分离细胞总RNA，然后从中纯化出mRNA。由于大多数真核生物mRNA分子的3′-端都有一段poly(A)尾巴，这种特殊的序列结构为从总RNA中分离纯化mRNA提供了方便。mRNA的分离方法中以寡聚（dT)-纤维素柱层析法最为有效，已成为常规方法。此法利用mRNA 3′-末端含有poly(A)的特点，在RNA流经寡聚（dT）纤维素柱时，在高盐缓冲液的作用下，mRNA被特异地结合在柱上，当逐渐降低盐的浓度时或在低盐溶液和蒸馏水的情况下，mRNA被洗脱，经过两次寡聚（dT）纤维柱后，即可得到较高纯度的mRNA。

2. cDNA第一链的合成

在cDNA文库的构建中，常采用两种方法合成cDNA第一链。一种方法是oligo(dT)引导的cDNA合成法，该法同样也利用了mRNA特殊的末端结构——poly(A)尾。以oligo(dT)为引物，在逆转录酶催化下合成与mRNA互补的cDNA第一链。有些大分子mRNA分子过长，用oligo(dT)作引物的逆转录难以到达mRNA的5′端，这是该方法的一个不足之处。为了解决这一问题，发展了另一种方法，即用随机引物引导的cDNA合成法，其原理是：根据许多可能的序列，合成出6～10个核苷酸组成的寡核苷酸片段（混合物），作为第一链合成的引物。在这种情况下，cDNA的合成可以从mRNA的许多位点同时开始。研究证实，用这种方法比较容易克隆得到过长的mRNA分子。

3. 双链cDNA的合成

（1）自我引导的合成法　合成了cDNA第一链以后，有多种方法合成与其互补的cDNA第二链。较早采用的方法是自我引导的合成法。该法是在合成了cDNA第一链后，用碱处理mRNA-cDNA杂交分子，使mRNA水解，解离的第一链cDNA 3′端回折形成发夹环(hairpin loop）结构，并以此作为合成cDNA第二链的引物合成双链cDNA，这样合成的双链cDNA带有发夹环结构，可用S1核酸酶切除（如图5-15所示），但这种切割作用会导致包括发夹环在内的部分cDNA序列信息的丢失，由于这个原因，现在已较少采用这种方法合成cDNA第二链。

（2）置换合成法　另一种合成cDNA第二链的方法是在mRNA-cDNA杂交分子形成后，用大肠杆菌RNase H降解杂交分子中的mRNA链，将其水解为许多短片段。这些短片段依然结合在cDNA第一链上，正好作为引物，在大肠杆菌聚合酶Ⅰ的作用下合成cDNA第二链。这样合成的结果是杂交链上的mRNA链完全被DNA片段取代，剩下的DNA片段之间的缺口由DNA连接酶连接，形成完整的cDNA第二链，仅在5′端留下一极小的空缺，如图5-16所示。使用RNase H合成cDNA第二链的方法，实际效果要比用S1核酸酶的方法强得多，因为它得到几乎完整的mRNA序列信息。

（3）同聚物加尾法　同聚物加尾法也用oligo(dT)作引物，在逆转录酶催化下合成cDNA第一链后，通过碱水解去掉RNA链，再经末端转移酶作用加上若干个单核苷酸，如

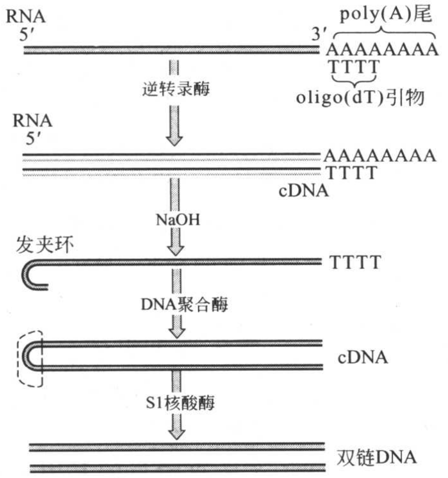


图 5-15　自我引导 cDNA 合成法示意图

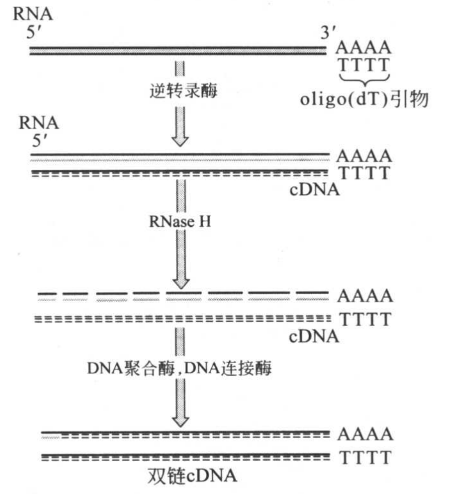


图 5-16　RNase H 降解法合成 cDNA 示意图

dG，随后使用一段合成的 oligo(dC) 作为引物在 DNA 聚合酶催化下合成 cDNA 的第二链，最后形成与 mRNA 相对应的完整双链 DNA(cDNA)，其一端为 oligo(dC)-oligo(dG) 双链片段，另一端为 oligo(dT)-oligo(dA) 双链片段，如图 5-17 所示。

4. 双链 cDNA 与载体连接构成重组体

在获得双链 cDNA 以后，可以通过几种不同的方法将其与载体连接构建重组体，并将重组体导入大肠杆菌宿主细胞，便可得到所需的 cDNA 文库。cDNA 与载体连接的常用方法有：cDNA 加接头连接法和直接以载体引导合成 cDNA 构建重组体的方法。

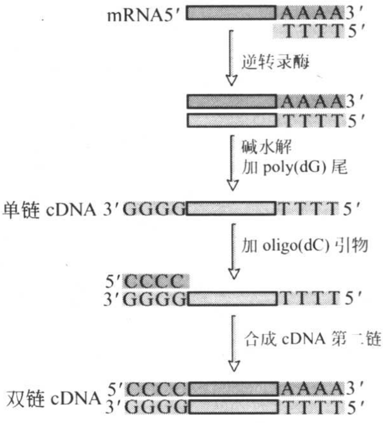


图 5-17　同聚物加尾法合成 cDNA 示意图

(1) cDNA 双链加接头连接法　cDNA 双链加接头连接的方法是在 cDNA 双链两端加上合适的接头 (linker)，实现载体与 cDNA 的连接。这种连接属于平末端之间的连接，需用 T4 DNA 连接酶催化，通过使用高浓度的接头分子，可以将接头分子与平末端的 cDNA 连接。由于使用了接头，因此在连接反应之前，需用特定的甲基化酶修饰 cDNA 双链，以避免 cDNA 在随后的酶切反应中被酶切消化。这样修饰过的 cDNA 分子，其两端接头中的特定限制酶位点，经相应限制酶消化后，便能产生黏性末端，可与经同样酶切消化的噬菌体载体互补连接，形成共价连接的重组分子。

为了获得全长的 cDNA 克隆，常采用同聚物加尾法获得的双链 cDNA 为起始材料，其优点是 cDNA 较为完整，当所获得的重组体与噬菌体衣壳蛋白进行外包装成为具有感染力的噬菌体颗粒后，感染大肠杆菌，在获得的大量噬菌斑中，每一个噬菌斑都是带有完整 cDNA 信息的单一的重组噬菌体，上述过程如图 5-18 所示。

(2) 载体引导合成 cDNA 构建重组体　利用载体自身的酶切位点及结构特点，通过同聚物加尾法合成 oligo(dT)，并以此为引物逆转录合成 cDNA 的第一链，在载体与 cDNA 连

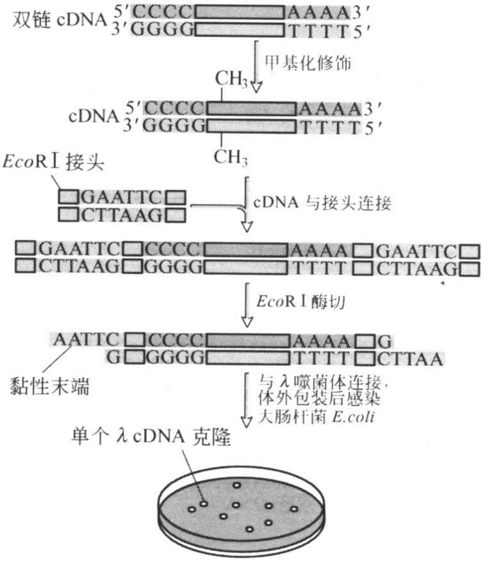


图 5-18 λ噬菌体 cDNA 文库的构建

接环化后，通过聚合作用合成 cDNA 的第二链，即构成重组体。构建过程如图 5-19 所示。

由图中可见，这种方法中间不需合成 cDNA 第二链，只在重组体建好后补上第二链。经验表明，用这种方法克隆全长 cDNA 的效率相当高，但同时由于步骤较多，操作有一定难度。

(3) 载体引导加衔接头连接构建 cDNA 重组体 这是一种十分有效的全长 cDNA 合成法，它同前一方法一样利用质粒载体引导 cDNA 的合成，并利用衔接头使线性重组体分子环化，形成双链环形重组体，即可有效转化大肠杆菌宿主细胞。反应过程见图 5-20 所示。

通过构建 cDNA 文库克隆目的基因具有突出的优势：①由于 cDNA 来源于 mRNA，而细胞在不同时间、不同条件下，其 mRNA 表达的种类和强度不尽相同，所以 cDNA 文库具有组织细胞特异性；因此，对于克隆筛选在这些

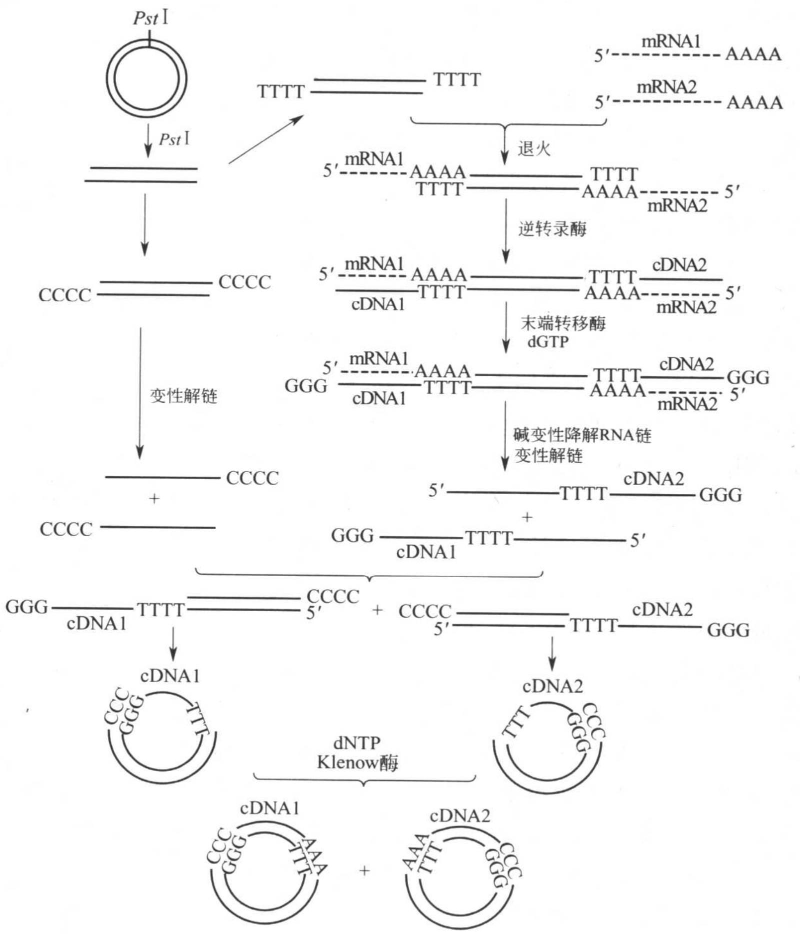


图 5-19 载体引导合成全长 cDNA

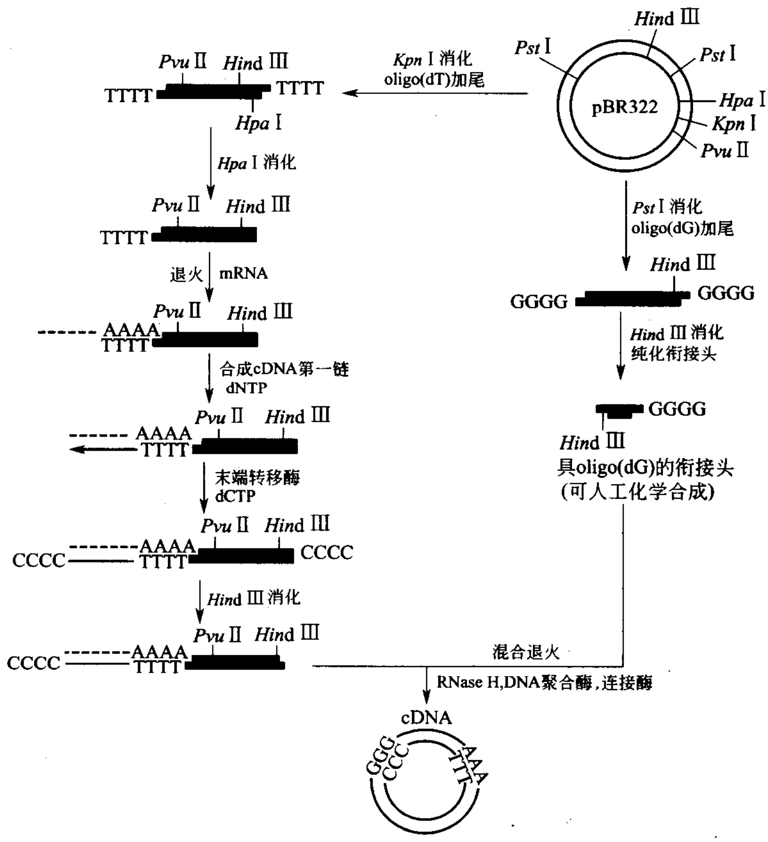


图 5-20　载体引导加衔接头连接构建 cDNA 重组体

组织中高表达的基因比较容易；②cDNA 不含内含子序列，可以直接决定蛋白质的氨基酸序列，因此 cDNA 文库要比基因组 DNA 文库小得多，能够比较容易从中筛选克隆得到细胞特异表达的基因。通过构建 cDNA 文库克隆目的基因的方法目前已经得到广泛地应用。

（二）构建基因组文库克隆目的基因

基因组文库（genomic library）也称基因文库（gene library），是含有某种生物体全部基因片段的重组 DNA 克隆群体，又称为 G-文库。基因组 DNA 的克隆与 cDNA 克隆不同，它的出发材料是基因组 DNA，而不是 mRNA。

构建基因组文库时，先将原核或真核细胞染色体 DNA 提纯，用物理方法（超声波、搅拌剪力等）或酶法（限制性核酸内切酶的不完全酶解）将 DNA 降解成预期大小的片段，然后将这些片段与适当的载体（常用噬菌体、黏粒或 YAC 载体）连接，转入受体细菌或细胞，这样每一个细胞接受了含有一个基因组 DNA 片段与载体连接的重组 DNA 分子，而且可以繁殖扩增，许多细胞一起组成一个含有基因组各 DNA 片段克隆的集合体，即构建了基因组 DNA 文库。如果这个文库足够大，能包含该生物基因组 DNA 的全部序列，就可以在建立基因文库后，再选用适当方法（如通过探针筛选法、免疫筛选法）从众多转化子菌落中筛选出含有目的基因的菌落，进行扩增、分离、回收，获得目的基因。当基因组比较小时，此法较易成功；当基因组很大时，构建完整的基因组文库本身就非易事，从庞大的文库中去克隆目的基因的工作量相当大。

在大多数 DNA 克隆操作中，一般喜欢采用质粒作为载体（如上述的 cDNA 克隆），因为其操作过程相对简单。然而，在构建基因组文库时，采用噬菌体载体则具有更多的优势。例如，取代型噬菌体载体的实用装载容量为 9～25kb，以每个克隆装载 20kb 片段计算，覆盖人类基因组全部 DNA（约 3×10^9 bp）所需克隆数约为 1.5×10^5 个，这对转染效率极高的噬菌体来说并不困难，而对转化效率相对较低的质粒则难以实现。以下将主要介绍以 λ 噬菌体为载体构建基因组文库的方法。

1. 基因组 DNA 的片段化

获得基因组 DNA 片段的方法有两种：一种是机械切割随机断裂 DNA，这种方法获得的 DNA 片段克隆效率低，不足以获得有代表性的文库；另一种是用限制酶对基因组 DNA 部分消化的方法，这种方法最早由 T. Maniatis 等（1982）提出，目前已被普遍采用。早期 T. Maniatis 等使用的方法是用两种限制酶混合消化基因组 DNA，他们所选用的两种限制酶都具有 4 个核苷酸的识别序列，故对于 DNA 有较高的切割频率。用双酶消化所产生的 DNA 限制片段的平均长度将小于 1kb。如果控制酶切消化条件，使其部分消化，就能得到平均 10～30kb 的随机分布的 DNA 片段。后来发展了单酶水解的方法，使操作简单方便，该法采用另一具有 4 个核苷酸识别序列的限制酶（*Sau*3AⅠ）单酶水解，因其产生黏性末端，并且是多种限制酶的同尾酶，易于与载体连接，因而更具优势。如果考虑用 λ 噬菌体作载体，则在单酶水解后分离其中的 20kb 左右的随机片段，这样大小的片段适合于取代插入 λ 噬菌体载体，经体外包装，便能回收到足够数量的独立的重组体，以构成一个完整的基因文库。

2. λ 噬菌体的体外包装构建重组噬菌体

作为克隆载体，要求 λ 噬菌体应能够进入溶菌周期，即能够合成头部和尾部蛋白并能裂解宿主菌，释放出 λ 噬菌体颗粒。因此在载体中必须保留与这些功能相对应的基因，其他功能的基因则可以被外源基因替换，这样在 λ 噬菌体中约可插入 25kb 的外源 DNA 片段。

在感染的宿主细胞内，λ 噬菌体的组装是分步进行的。首先，λ 噬菌体分别组装其头部和尾部。头部外壳并不直接与 λ DNA 包装，而需要两种特定的 λ 噬菌体编码的蛋白质辅助。在 λ DNA 在复制过程中形成的 DNA 分子可以借助其 *cos* 末端相互连接，形成长的多聚 DNA 分子，称为多连体（concatomers），它含有多份的 λ 基因组成分。包装时 λ 噬菌体编码的 Nu1 蛋白质和 A 蛋白质可以结合 *cos* 位点，并将两个相邻 *cos* 位点间的 DNA 装入 λ 噬菌体头部结构中，这种组装方式导致所形成的头部只含一份 λ 基因组。一旦 λ DNA 进入头部就可与预先合成的 λ 噬菌体尾部结合，形成完整的噬菌体颗粒，如图 5-21 所示。

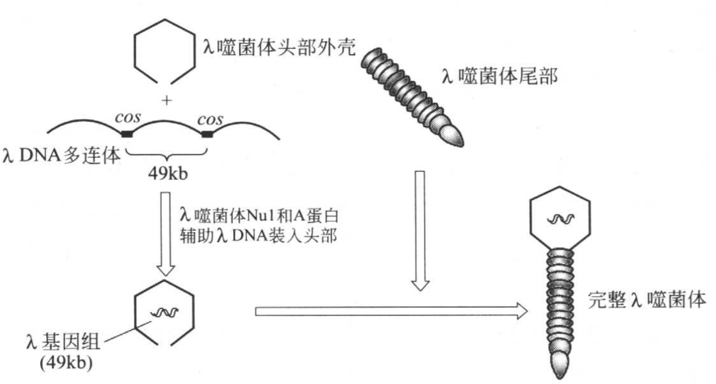


图 5-21　λ 噬菌体的组装

含外源 DNA 片段的重组 λ 噬菌体的制备可以采用取代的方法，即用外源 DNA 片段取代 λ 噬菌体的重组区构成重组体，在体外与预先提取的头部和尾部结构包装成完整的噬菌体颗粒。

3. 构建基因组文库的两种方案

（1）*Hae*Ⅲ和 *Alu*Ⅰ双酶部分水解加接头构建基因组文库　双酶部分消化常选用两种识别序列不同的限制酶 *Hae*Ⅲ（5′…GG↓CC…3′）和 *Alu*Ⅰ（5′…AG↓CT…3′）。然而，这两种限制酶切割形成的都是平末端的 DNA 片段分子，所以在与载体分子重组之前需要加用适宜的接头，例如 *Eco*RⅠ接头，经 *Eco*RⅠ酶切之后即成为特定的黏性末端，可以与同样经过 *Eco*RⅠ限制酶消化处理的取代型 λ 噬菌体载体，通过黏性末端间的互补作用有效地连接（但应注意在用 *Eco*RⅠ酶切消化接头之前需用甲基化酶作必要的处理，以保护 DNA 片段内部的 *Eco*RⅠ位点）。经 DNA 连接酶连接封口并在体外包装之后，便可有效地导入大肠杆菌宿主细胞，获得所需的基因组文库，如图 5-22 所示。

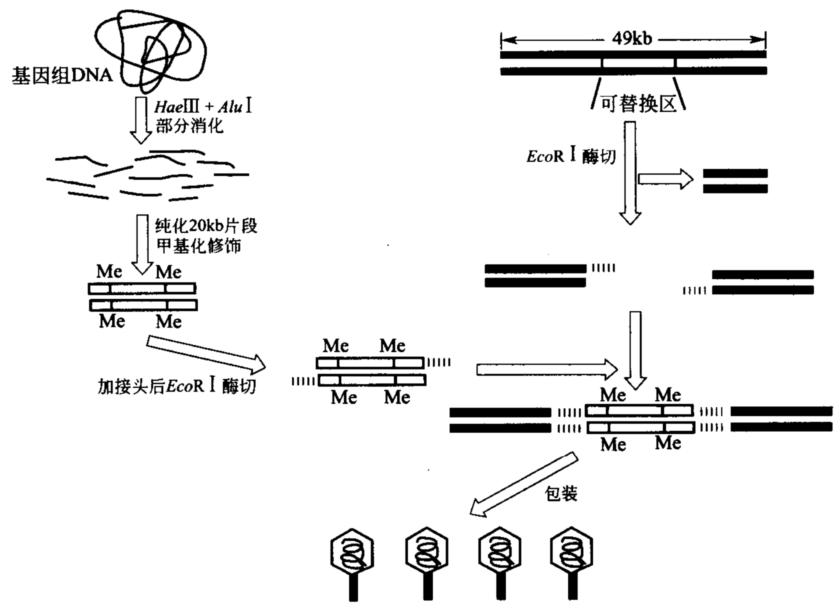


图 5-22　通过接头连接法在 λ 噬菌体载体上构建基因组文库

（2）*Sau*3AⅠ单酶水解构建基因组文库　单酶消化使用限制酶 *Sau*3AⅠ，该酶的识别序列亦为 4 个核苷酸，但是水解产生黏性末端。*Sau*3AⅠ部分消化能够较为方便地获得随机的 DNA 限制片段并由于它与 *Bam*HⅠ是一对同尾酶，都形成 GATC 黏性末端，因此由 *Sau*3AⅠ酶切消化产生的 DNA 片段可容易地插入到经 *Bam*HⅠ消化过的 λ 噬菌体载体上。这类局部消化法与 λ 噬菌体载体的体外包装技术相结合，已成为当今最广泛使用的构建真核基因组文库的实验方案之一。它的基本程序如图 5-23 所示。首先，提取的真核基因组 DNA，用限制酶 *Sau*3AⅠ部分消化；然后，将消化产物通过琼脂糖凝胶电泳或蔗糖梯度离心，分部收集 15～20kb 范围的 DNA 片段。与此同时，λDNA 则用 *Bam*HⅠ酶切，分离纯化的酶切产物中的 λDNA 的两臂分子，在 T4 DNA 连接酶作用下与上述收集的基因组 DNA 片段连接。连接产物在体外包装体系中组装成具感染力的噬菌体颗粒，感染大肠杆菌细胞后即可获得重组体克隆库，即基因组文库。它们包含有真核生物（如哺乳动物）基因组的

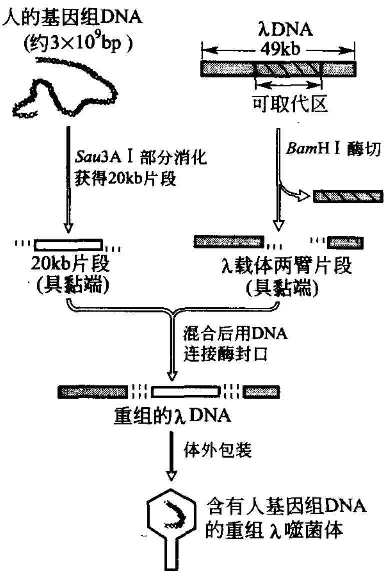


图 5-23 *Sau*3A Ⅰ部分消化法构建基因组文库

序列。

4. 应用柯斯质粒载体构建基因组文库

λ噬菌体作载体构建基因组文库已经成为基因组文库构建的常用方法。除此以外，柯斯质粒载体也可用于基因组文库的构建。由于柯斯质粒载体可携带长度约为45kb的外源DNA，并能在体外被高效地包装，因此适合于较大基因片段的克隆筛选。在以柯斯质粒作载体的文库构建中，基因组DNA可用*Sau*3AⅠ部分消化，回收35～45kb的片段群体，与*Bgl*Ⅱ消化的柯斯质粒载体DNA连接重组。经体外包装之后，感染大肠杆菌宿主细胞。在细胞内，柯斯质粒载体按质粒特性进行复制扩增，获得基因组文库。

四、利用差异文库的克隆方法

（一）差别杂交和扣除杂交克隆的目的基因

差别杂交（differential hybridization）适用于分离经特殊处理而被诱发表达的mRNA之cDNA克隆。并在此基础上发展出了扣除杂交（subtractive hybridization）技术。扣除杂交又叫扣除cDNA克隆（subtractive cDNA cloning），它是通过构建扣除文库得以实现的。

1. 差别杂交

差别杂交的技术基础十分简单，它不需要任何有关的目的基因的核苷酸序列信息，而重要的是要拥有两种不同的细胞群体：在一个细胞群体中目的基因正常表达，在另一个细胞群体中目的基因不表达。因此，可以通过这两种细胞总mRNA（或是它们的cDNA拷贝）制备两套探针。两套探针分别与同一菌落克隆平板的转印膜杂交，在X射线胶片上曝光后，比较这两种底片并对照原平板，便可以挑选出含目的基因的菌落，如图5-24所示。

由于差别杂交的灵敏度较低、筛选杂交滤膜的工作量大以及重复性差等缺点，使其应用受到很大限制，尤其对低丰度mRNA显得更加突出。为此，进一步发展了扣除杂交克隆方法。

2. 扣除杂交

扣除杂交法的本质是除去那些普遍共同存在的cDNA序列，从而使欲分离的目的基因的序列得到有效的富集，提高了分离的敏感性。

例如，某基因只在A细胞中表达，而不在B细胞中表达。那么从A细胞mRNA制备的单链cDNA，与过量的B细胞来源的mRNA杂交，在有利于DNA-RNA杂交的条件下，将会产生3种主要产物：①DNA-RNA杂合双链，它们由A和B两类细胞中同时表达的该基因mRNA/cDNA组成；②cDNA单链，它来自于A细胞中特有的基因表达的mRNA；③杂交后剩余的mRNA链。当上述杂交混合物通过羟基磷灰石柱时，DNA-RNA杂交分子便结合在柱上，而游离的单链cDNA和RNA则过柱流出。降解RNA后即可获得A细胞特有的cDNA，经转变为双链cDNA之后，与适当的λ噬菌体载体重组并转化大肠杆菌寄主细胞，这样便得到了A细胞特有的高度富集的cDNA扣除文库。然后，再按照同样方法制备扣除的cDNA探针，筛选文库即可获得A细胞特异表达的目的基因，如图5-25所示。

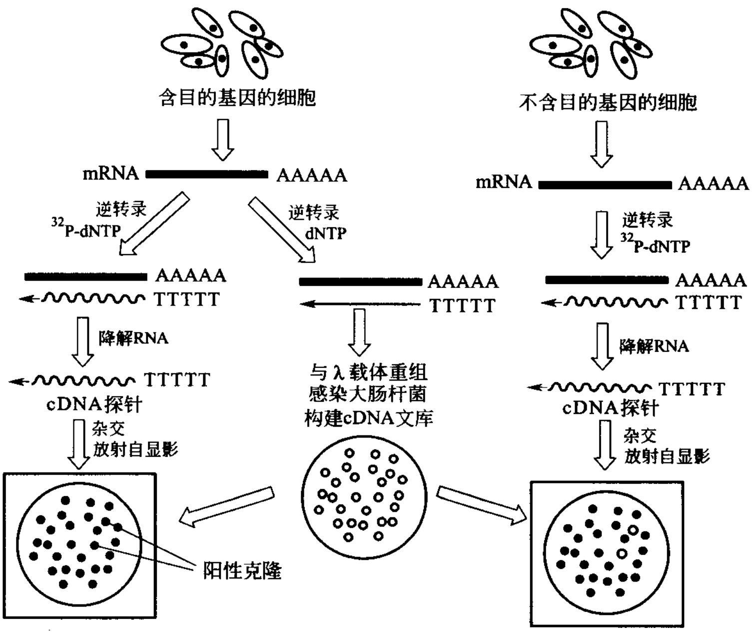


图 5-24 差别杂交克隆目的基因

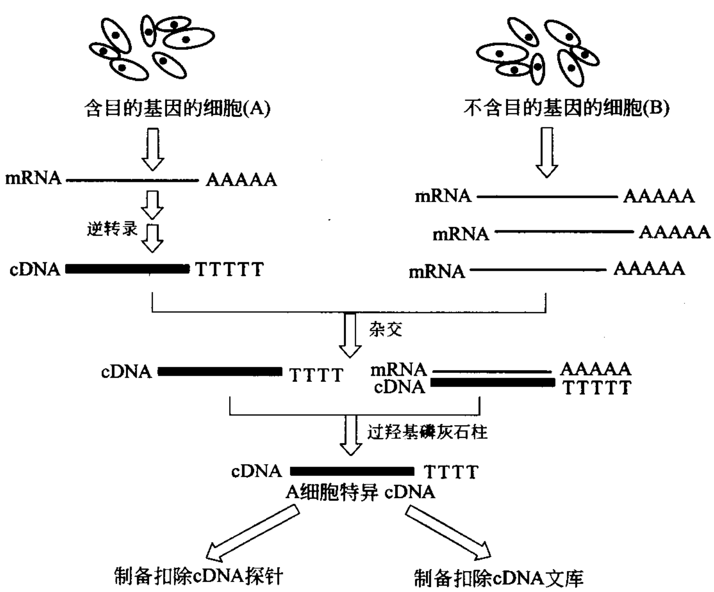


图 5-25 扣除杂交克隆目的基因

扣除杂交法利用杂交除去共有的 mRNA，减少了构建文库的 cDNA 种类，使含目的基因的 cDNA 得以有效地富集，因而极大地减少了克隆的工作量，同时提高了克隆的效率和敏感度。这一方法目前有了进一步的改进，并发展了新的操作技术，如使用磁珠的扣除技术(magnet assisted subtractive technique，MASP)、oligo(dT) 乳胶结合 PCR 的方法等。

（二）代表性差别分析技术克隆目的基因

代表性差别分析（representation difference analysis，RDA）技术最早是由美国冷泉港

实验室的 N. Lisitsyn 和 M. Wigler 于 1993 年提出，用来检测两个个体之间 DNA 多态性的差异。M. Hubank 等应用 RDA 克隆差异表达基因，创建了 cDNA 代表性差异分析法，即 cDNA-RDA。cDNA-RDA 技术的基本原理是将 PCR 与扣除杂交相结合，通过降低 cDNA 群体的复杂度和多次更换 cDNA 两端接头引物等方法，分离克隆两个复杂 cDNA 群体中差异表达的基因。目前，RDA 技术已被广泛地应用于分离编码产物未知的基因。

1. 基本原理

代表性差别分析（RDA）中将两种差异对象来源的 DNA 分别称为 tester DNA（检测 DNA）和 driver DNA（驱动 DNA），前者含待检测的基因。cDNA-RDA 的主要步骤如下。

（1）确定分析体系，获得 cDNA 群体　选择两个用于 RDA 分析的体系，提取总 RNA，采用与 DDRT-PCR 相似的方法合成双链 cDNA，即可获得两种 cDNA 群体。

（2）制备检测 DNA 和驱动 DNA　用识别 4 个碱基的限制性核酸内切酶分别将两种 cDNA 群体酶切消化成带黏性末端的片段，再连上含有该酶酶切位点的寡核苷酸连接物，PCR 扩增后用同一内切酶切下连接物，此时对检测组 DNA 应再连上另一种含有该酶酶切位点的连接物制成检测 DNA；而驱动组则不需要连上另一种接头，因此酶切后即已制成驱动 DNA。这样检测 DNA 和驱动 DNA 各自代表了两种细胞 cDNA 的一部分，且由一系列小片段 DNA 组成，降低了 cDNA 的复杂度，使后续杂交步骤快速、有效。

（3）液相杂交　检测 DNA 和驱动 DNA 按一定的比例（1/100～1/200），在适当的缓冲体系中高温变性，中温（67℃）长时间复性（20h）后，90%的单链 DNA 复性成稳定的双链结构，出现 4 种情况的组合 DNA 片段：检测 DNA 中仅有的片段自身复性；检测 DNA 和驱动 DNA 复性形成异源双链；驱动 DNA 自身复性及检测 DNA 和驱动 DNA 都剩一些单链。

（4）特异片段的富集　在上述 4 种组合的 DNA 片段中，经核酸酶处理除去单链 DNA，这时只有检测 DNA 两端带有接头；将其两端填平后，以填平的末端为引物的结合点，以原接头为引物进行 PCR 扩增，扩增产物正是检测 DNA 有而驱动 DNA 中缺失的 DNA 片段。以此 PCR 产物作为检测 DNA 再次与过量的驱动 DNA 杂交，重复三轮杂交后即能充分富集出检测 DNA 表达而驱动 DNA 不表达的 cDNA 片段，富集倍数可达 10^6 倍。上述过程如图 5-26 所示。

2. cDNA-RDA 的特点

cDNA-RDA 综合了 DDRT-PCR 和 RDA 两种技术的特点，因此也具备这两项技术各自的优点，并在一定程度上克服了它们的弊端。首先，cDNA-RDA 能发现与表型相关的基因异常，并能随基因表达的时效性不同而分离出表达差异的基因。其次，由于 cDNA 是从 mRNA 逆转录而来，片段长度有限，加上应用识别 4 个碱基的限制性核酸内切酶进行消化，使得基因组的复杂程度得到了极大的简化，提高了基因克隆效率。利用 cDNA-RDA 所获得的 EST（表达序列标签）可以较容易地从 cDNA 文库中筛选出基因的全长序列。

然而 cDNA-RDA 也存在一些问题：①对 mRNA 要求较高。由于 cDNA 从 mRNA 逆转录而来，因而要求样品必须新鲜，其 RNA 要有一定的质量保证，导致 mRNA 降解的任何因素都可能造成信息的大量损失；②可能存在漏检。与基因组 DNA 相比，cDNA 片段要短得多，由此造成限制性核酸内切酶酶切位点分布不均，其消化的结果可能造成 cDNA 序列两端信息量丢失，可能使得在酶切位点区域外的点突变、小的缺失或插入难以被发现；或完全无酶切位点的片段而无法被 RDA 检测到；③操作较复杂。RDA 在方法上同样存在繁琐、

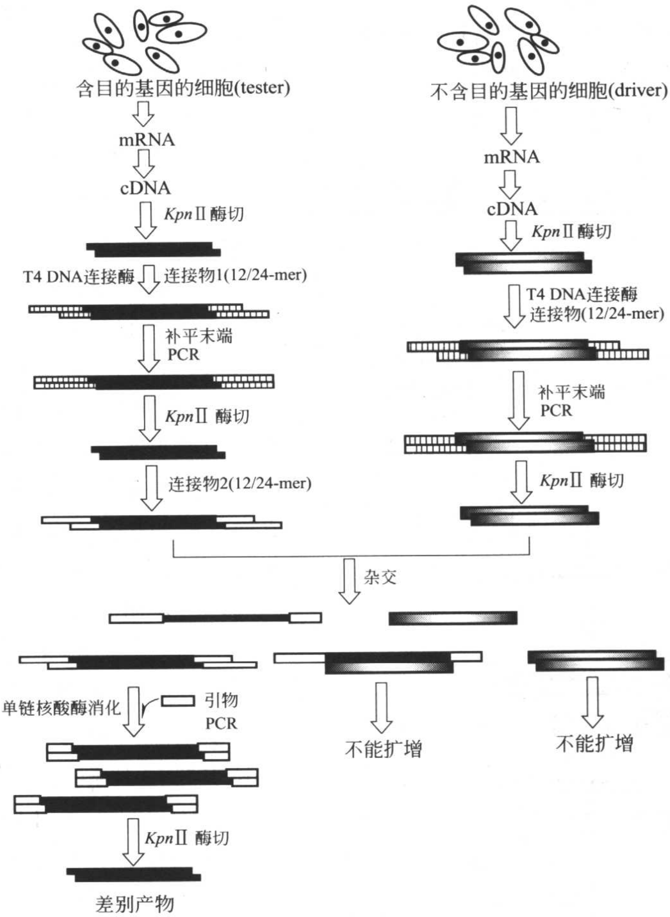


图 5-26 cDNA-RDA 反应步骤示意图

费时、费力和应用过程中易于污染等缺点。

（三）抑制性扣除杂交技术

抑制性扣除杂交（supression subtractive hybridization，SSH）技术是一种基因克隆的新方法，它是由 L. Diatchebko 于 1996 年在 RDA 的基础上建立起来，主要用于研究细胞生殖、发育、分化、癌变、衰老及程序化死亡等生命过程中有关基因的差异表达以及差异表达相关基因的克隆。SSH 是 RDA 技术的发展，能有效克服 RDA 或 cDNA-RDA 技术难以解决的问题，如两者不能用于分离两组基因表达差异较小的基因，也不能用于研究存在上调表达的基因等。

1. SSH 的主要原理

SSH 技术是以抑制 PCR（suppression PCR）为基础的 DNA 扣除杂交方法，它将检测 DNA（tester DNA）单链标准化步骤和消减杂交步骤结合为一体。其中标准化步骤均等了检测子 cDNA 单链丰度，而消减杂交步骤去除了检测子（tester）和驱动子（driver）之间的共同序列，使检测子和驱动子之间不同的序列得到扩增。因此 SSH 显著增加了获得低丰度表达差异 cDNA 的概率，简化了对消减文库的分析。

抑制 PCR 是利用链内复性优先于链间复性的原理，使非目标序列片段两端的长反向重

复序列（long inverted repeats）在复性时产生“锅柄样”（panhandle-like）结构而无法与引物配对，从而选择性地抑制了非目标序列的扩增。同对，根据杂交的二级动力学原理，丰度高的单链 cDNA 复性时产生同源杂交速度要快于丰度低的单链 cDNA，从而使得丰度存在差异的 cDNA 相对含量趋于基本一致。

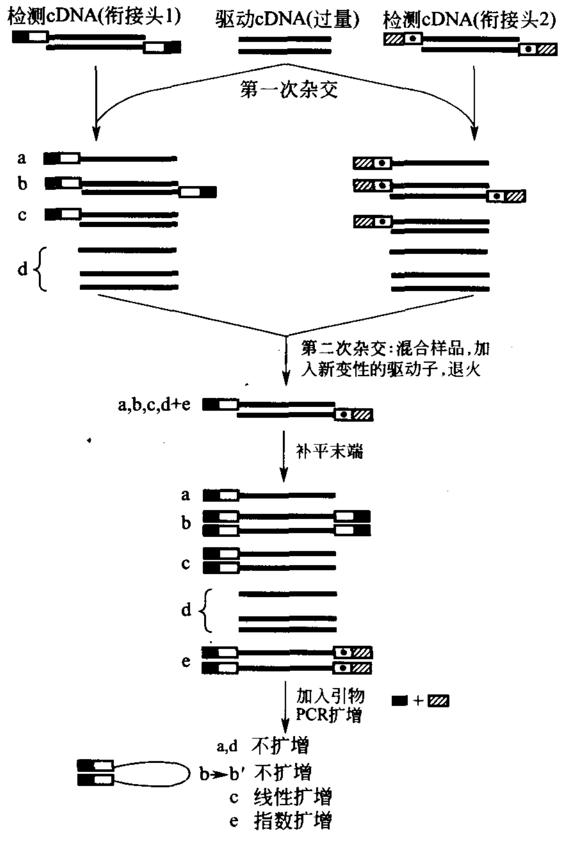


图 5-27　SSH 方法示意图

2. SSH 的基本过程

SSH 的主要步骤如下，如图 5-27 所示。

（1）制备检测 cDNA 和驱动 cDNA　制备两种试验材料的 mRNA 并逆转录成 cDNA，分别作为检测 cDNA 和驱动 cDNA。用限制性核酸内切酶切割，产生大小适当的平头末端 cDNA 片段。

（2）与衔接头连接　将检测 cDNA 分成均等的两份，分别接上衔接头 1 或衔接头 2，此衔接头由一长链（40nt）和一短链（10nt）组成的一端是平末端的双链 cDNA 分子。长链 3′-端与 cDNA 5′-端相连。长链外侧序列（约 20nt）与第一次 PCR 引物序列相同，内侧序列则与第二次 PCR 引物序列相同。此外，接头上含有 T7 启动子序列及内切酶识别位点（如 *Not* Ⅰ、*Srf* Ⅰ、*Sma* Ⅰ 和 *Xba* Ⅰ 等），为以后将该片段插入克隆载体和测序提供便利。

（3）第一次差减杂交　用过量的驱动 cDNA 分别与上述两份检测 cDNA 混合，变性后复性杂交，产生 4 种产物：a. 是单链检测 cDNA；b. 是自身复性杂交的检测 cDNA 同源双链；c. 是检测 cDNA 和驱动 cDNA 的异源双链；d. 是驱动 cDNA（单链和双链）。第一次杂交的目的是实现检测单链 cDNA 的标准化（normalization），也就是使原来有丰度差别的单链 cDNA 的相对含量达到基本一致。由于检测 cDNA 中与驱动 cDNA 序列相似的片段大都和驱动 cDNA 形成异源双链分子（c），使检测 cDNA 中的差异表达基因的目标 cDNA 得到大量富集，这是第一次杂交后的第二个目的。

（4）第二次差减杂交　合并第一次杂交后的两份杂交产物，再与过量的驱动 cDNA 进行变性、复性杂交，此时，只有第一次杂交后经丰度均等化和消减的单链检测 cDNA 和驱动 cDNA 一起形成各种双链分子，这次杂交进一步富集了差异表达基因的 cDNA，并产生了一种新的双链分子（e），它的两个 5′-端分别连上了来自两个样本的不同的接头。

（5）第一次 PCR 反应　在上述产物中加入一对分别与衔接头 1、衔接头 2 外侧端序列相同的引物（约 20～22 个核苷酸）作 PCR 扩增，这时只有两端是不同接头的双链 cDNA 分子（e）才能呈指数扩增，而两端连上相同接头的同一片段（b）由于末端保留了长的反向重

复序列，在变性复性过程中形成“锅柄样”结构而不能呈指数扩增。因此，最终的结果是驱动 cDNA 与检测 cDNA 中的相同部分得到抑制，而在驱动 cDNA 中缺乏但却存在于检测 cDNA 的序列，即差异表达 cDNA 序列得到显著的扩增。

(6) 第二次 PCR 反应　选用一对分别与接头内侧端序列相同的引物进行第二次 PCR，基于 PCR 抑制效应的存在，可以特异性扩增代表了差别表达的 cDNA 片段。

(7) 差异片段的初步筛选　经上述 PCR 所得的代表性差别表达的 cDNA 片段可以利用接头上存在的酶切位点，插入适当载体中，然后转化细菌。经 X-gal 蓝白斑初步筛选出具有插入的克隆，再经探针杂交找出具有差别表达的 cDNA 片段。然后，对这些 cDNA 片段进行序列测定、序列同源性分析等一系列工作，完成基因的克隆、鉴定。

3. SSH 技术的特点

SSH 方法的假阳性率低，这是该方法的最大优点。Stein 等的研究发现，SSH 的阳性率可达 94%，这是由它的两步杂交和两次 PCR 所保证的。另外，SSH 的敏感性高、速度快、效率高。通常，一次 SSH 反应可以同时分离几十或成百个差异表达基因。不足之处主要有：需要较多量的 mRNA（数微克），如果量不够，那么低丰度的差异表达基因的 cDNA 很可能会检测不到，这是 SSH 技术不能被广泛应用的主要障碍；另外，获得的 cDNA 片段长度有限。

综上所述，SSH 由于操作简便，亦无需昂贵的实验仪器，故逐渐成为目前寻找有差别表达的基因的适用方法。睾丸和卵巢特异表达基因以及乳腺癌雌激素受体某些相关基因等的分离克隆都是成功应用 SSH 技术完成的。

第五节　克隆筛选与鉴定

外源基因与载体连接形成重组 DNA，重组 DNA 分子虽然具备复制或表达的潜力，但必须在宿主细菌（细胞）内才能实现复制与表达。另外，所构建的重组体是否正确也需要验证，因此，在构建了重组体以后，接下来的工作就是将重组 DNA 转入受体细胞，并对转化菌进行筛选和鉴定，以获得大量含目的基因的阳性重组子，进一步研究该基因的结构与功能。

一、重组 DNA 导入受体细胞

重组 DNA 分子导入特定的受体细胞（宿主细胞），使之获得新的遗传表型，这个过程称为重组 DNA 分子的转化（transformation）。转化现象在自然界普遍存在，是常见的基因转移方式之一。然而，在自然界发生的转化与 DNA 重组技术中使用的转化在原理和方式上均显著不同。在 DNA 重组技术中通过使用各种人工方法将外源基因导入受体细胞，在体外实现转化。不同宿主细胞转化的方法各不相同。细菌尤其是大肠杆菌的常用的转化方法主要是 $CaCl_2$ 转化方法和电穿孔转化方法等。

重组体转化过程中应注意选择合适的受体细胞，并应注意转化效率和生物安全性。一般，野生型细菌对外源 DNA 的转化效率较低，并且有可能感染其他生物种群，因此不能作

为基因工程的受体菌。目前使用的受体菌必须是经过遗传改造的安全、高效的菌株，这种受体菌具有以下特点：①限制修饰系统缺陷，如 *hsd*R⁻ 遗传表型的大肠杆菌菌株，增加外源 DNA 的转化效率；②重组系统缺陷，如选择 *rec*A 缺陷（*rec*A⁻）型大肠杆菌细胞可使重组频率比正常降低 10^6 倍；③易于转化，如大肠杆菌膜蛋白 LamB 是 λ 噬菌体吸附所必需的结构；④可筛选的遗传表型，如对抗生素的抗性及营养缺陷表型等；⑤感染寄生缺陷，即不会因受体菌的感染寄生作用而广泛传播。

受体细胞的上述特点对外源基因的克隆与表达来说都是十分重要的，它直接关系到基因工程的成败。对于以改良生物物种为目的的克隆工作而言，受体细胞的选择则没有太大的选择余地，此时的受体细胞只能是待改良的生物物种本身。

二、重组体的筛选与鉴定

重组体导入受体细胞以后，经过短时间的培养扩增即可用于进一步的筛选与鉴定，这是基因克隆的最后一个环节。所谓筛选（screening）是指通过某些特定的方法，从被转化的细胞群体或基因文库中，鉴别出真正具有重组 DNA 分子的特定克隆（即阳性重组子）的过程。筛选出来的阳性重组子通过培养扩增，以获得所需基因片段的大量拷贝，并通过进一步鉴定，最终确定其正确性，为后续研究奠定基础。

不同的克隆载体及相应的宿主系统，其重组子的筛选、鉴定方法不尽相同，概括起来主要有以下几类。

（一）根据遗传表型的筛选方法

重组体分子一般带有遗传标记，它赋予转化细胞特定的遗传表型。遗传表型筛选的原理就是根据转化细胞与非转化细胞遗传表型的差异，通过在培养基中添加一些相应的筛选物质，使不同表型的细胞区别开来，从而获得具有同一种表型的转化细胞克隆。这种筛选方法操作非常简单，常作为重组体筛选的第一步。

1. 抗生素抗性筛选

这是最常用的转化细胞筛选方法。因为大多数基因工程载体均带有抗生素抗性基因，如氨苄青霉素抗性基因、四环素抗性基因、卡那霉素抗性基因等。用此类载体构建重组体时，如果保留了抗性基因的活性，那么它所转化的细胞可以在含此种抗生素的培养基中生长出转化子克隆，而未转化的细胞则不能生长，这样可以初步筛选出转化子克隆，这是一种正向选择方式。以常见的 pGEM-3Z 质粒载体为例，该载体含有氨苄青霉素抗性基因（amp^r）的选择标记，外源基因通过多克隆位点插入载体分子，重组体转化宿主菌 JM109，在含氨苄青霉素的培养基中能够生长形成克隆的即是阳性转化子，未转化的宿主菌在这种培养基中不能生长。应当注意，除了阳性重组子之外，自身环化的载体、未酶解完全的载体以及非目的基因插入载体形成的重组子均能转化细胞并形成菌落，故本法仅是阳性重组子的初步筛选。

2. 插入失活筛选

在上述利用抗生素抗性筛选得到的阳性克隆中，除了重组子以外，还含有空载体等其他非重组体的转化子。为进一步筛选出重组子，可以利用插入失活方法将重组子与非重组子分开。使用插入失活筛选方法需要载体具有两种抗生素的抗性基因，如 pBR322，它含有氨苄青霉素和四环素的抗性基因，当外源基因插入到其中一个抗性基因位点（如 amp^r 中的位点）时，氨苄青霉素抗性基因被外源基因所分割而失去活性，但其四环素的抗性基因保持活性，这种重组体转化的大肠杆菌能够在含有四环素的培养基中生长，但不能在含有氨苄青霉

素的培养基中生长，而空载体或两个抗性基因保持完好的非期望重组子均可以在这两种培养基中生长，这样就可以将期望的重组子筛选出来。

3. 插入表达筛选

插入表达筛选与插入失活筛选正好相反，它利用外源基因插入特定的调控基因序列中，导致调控基因失活而不能产生出活性的阻遏蛋白，这样便激活了相应标志基因的表达。例如pTR262 质粒载体就是用于插入表达筛选的特定载体，该载体由 pBR322 衍生而来，在其四环素抗性基因的上游含有一段 λ 噬菌体 DNA 的 *cI* 基因及其调控序列，*cI* 基因编码一种阻遏蛋白，可以抑制四环素基因的表达。*cI* 中含有外源基因的插入位点，当外源基因在此插入时，*cI* 基因失活，不能产生有活性的阻遏蛋白，对四环素抗性基因的抑制作用解除。因此，含有重组体的转化菌可以在含四环素的培养基中生长，形成菌落，如图 5-28 所示。

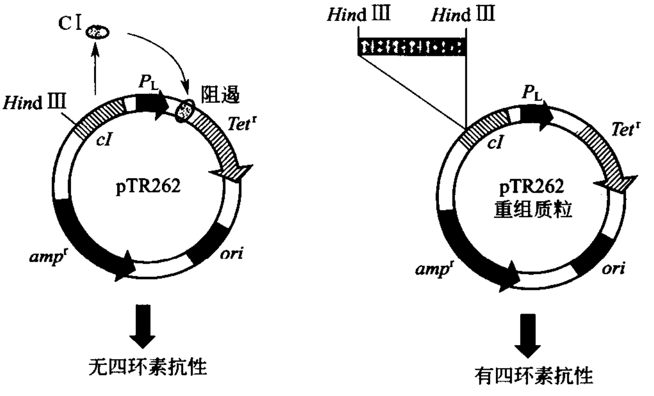


图 5-28 四环素抗性正选择系统

4. *β*-半乳糖苷酶系统筛选

用于这种筛选系统的质粒载体含有 *lac*Z′标记基因，它由两部分组成，即大肠杆菌 *β*-半乳糖苷酶基因的调控序列和该酶氨基端 146 个氨基酸组成的 *α*-肽的编码序列，其表达产物为无活性的酶的一部分（*α*-肽）。此类质粒载体的宿主菌在其染色体 DNA 上含有 *β*-半乳糖苷酶另一半的编码序列（即该酶羧基端的部分序列），其表达产物也是无活性的片段。当质粒表达的氨基端片段和宿主菌表达的羧基端片段同时存在于一个体系时，无论是在细胞内还是在细胞外，两片段均能结合，产生出具有完整酶活性的 *β*-半乳糖苷酶，可以将无色的作用底物 X-gal（5-溴-4-氯-3-吲哚-D-半乳糖苷）水解为蓝色产物，上述反应过程即所谓 *α*-互补作用。

外源基因可以利用载体 *lac*Z′中的多克隆位点插入载体而构建成重组体。然而这种插入作用导致 *lac*Z′基因失活，从而破坏了 *α*-互补作用，也就不能产生蓝色菌落，而产生白色菌落。有些载体还带有 *lac*Z 的抑制基因 *lac* Ⅰ，它编码的阻遏蛋白可以阻止 *lac*Z 的转录。在有诱导物 IPTG（异丙基-*β*-D-硫代半乳糖苷）或乳糖存在时，通过 IPTG 与阻遏蛋白结合，诱导 *lac*Z 的表达，因此在筛选时，培养基中需同时存在 X-gal 和 IPTG。

5. 根据插入序列表型特征的直接筛选

上述筛选都是利用重组体中载体的筛选标记进行的筛选，除此以外也可以利用目的基因的表达特征及其与宿主细胞的互补或抑制关系筛选重组子。在利用目的基因表达特征的筛选中，要求重组 DNA 能够在所转化的宿主菌中表达出外源基因的功能性蛋白产物，并使宿主菌获得新的表型特征，以区别非重组子。因此，能够用这种方法筛选的基因有一定限制，即

待筛选的目的基因要有合适的突变宿主菌或缺陷宿主菌，以产生互补效应或抑制效应。理论上，这种筛选方法是一种较为理想的直接筛选，但实际上符合上述条件的基因的数目有限，并受到表达效率的影响。目前，已经详细研究了许多种宿主菌（常用的如大肠杆菌）不同基因缺陷的突变株，这些突变菌株在克隆与宿主菌具互补效应的基因中发挥了重要作用，甚至有些真核基因也能成功地应用互补效应直接筛选获得；典型的代表如酵母细胞 *his* 基因的克隆，它是通过将酵母的 DNA 片段构建成重组体后转化大肠杆菌 *his*B 突变株，该菌株缺乏利用组氨酸的基因，通过营养筛选获得了酵母细胞 *his* 基因的克隆。

（二）依赖重组子结构特征的筛选法

上述根据遗传表型筛选出来的阳性克隆，还必须经过进一步鉴定，以获得含有目的基因的重组体。这一阶段的筛选与鉴定工作主要围绕重组体的结构特征进行分析鉴定，其中采用的方法包括酶切图谱分析、杂交分析、序列测定等。

1. 快速裂解菌落鉴定分子大小

从平板中直接挑取菌落，液体培养基扩增后裂解细菌，提取其中的 DNA，不需要限制性核酸内切酶消化，直接进行 DNA 凝胶电泳，与载体 DNA 比较迁移率，初步判断是否有插入片段存在及插入片段的大小。本方法适用于插入片段较大的重组子初步筛选。

2. 限制性核酸内切酶酶切图谱鉴定

上述分子大小的鉴定方法虽然可以初步判别是否含有重组体分子，但却难以判断重组体中是否含有期望的目的基因及目的基因的插入方向，因为在重组过程中，载体 DNA 分子可以与多种带相同末端的 DNA 分子连接构成重组体。要正确区分期望重组体和非期望重组体就要进行限制酶酶切图谱分析，具体做法是：将初步筛选鉴定的具有插入片段的重组子菌落进行小量扩增，再分离出重组质粒或重组噬菌体 DNA，用一种或多种相应的限制酶消化，通过 DNA 凝胶电泳分析确定其是否含有外源基因以及外源基因的插入方向。例如，某大小为 1200bp 的目的基因片段通过 *Bam*H Ⅰ 和 *Nhe* Ⅰ 双酶切后装入载体，如图 5-29 所示。在克隆鉴定时，可采用多种不同的酶切方式观察 DNA 片段大小：经 *Apa* Ⅰ 完全酶切消化后产生两个片段，大小分别为 6340bp 和 1970bp；用 *Bam*H Ⅰ 和 *Nhe* Ⅰ 双酶切消化，则也产生两个片段，其大小分别为 1200bp 和 7010bp；用 *Bam*H Ⅰ 和 *Apa* Ⅰ 双酶切消化，则产生三个片段，大小依次为 5360bp、1970bp 和 980bp。上述酶切消化片段均是在酶切消化反应完全充分的条件下的产物，如果酶切不完全，则会产生其他片段。

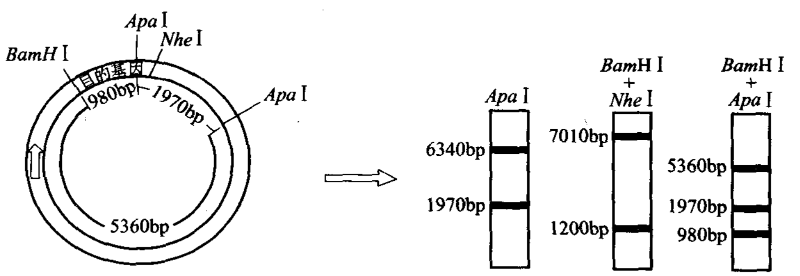


图 5-29 酶切鉴定图谱

3. 核酸分子杂交筛选

为了进一步确定 DNA 插入片段的正确性，在限制酶消化重组子并经凝胶电泳分离后，通过 Southern 印迹转移，将 DNA 转移到硝酸纤维素膜上，再用放射性同位素或非放射性标

记的相应外源 DNA 作为探针，进行分子杂交，鉴定重组子中的插入片段是否是所需的靶基因片段。

4. PCR 方法筛选重组子

由于构建重组体的载体分子的序列都是已知的，因此要确定外源基因是否插入载体分子或插入序列是否正确就可以采用 PCR 的方法进行鉴定。如 pGEM 载体系列中多克隆位点的两侧分别是 SP6 及 T7 启动子的序列，据此可以设计合适的引物，对小量抽提的质粒 DNA 进行 PCR 分析。这种方法不但可以迅速得到大量插入片段，而且能对扩增片段进行各种分析，包括 DNA 序列分析，从而获得插入片段正确性的最终结果。对大多数克隆工作来说，准确可靠的外源基因 DNA 序列测定结果是决定克隆成败的主要依据。

5. 菌落或噬菌斑原位杂交

菌落或噬菌斑原位杂交技术是至今为止最为通用的重组子筛选技术，它是先将转化菌直接铺在硝酸纤维素膜或琼脂平板上，再转移至另一硝酸纤维素膜上，用放射性同位素标记的特异 DNA 或 RNA 探针进行分子杂交，然后挑选阳性克隆菌落。本方法能进行大规模操作，对于从基因文库中筛选目的重组子，是一个首选的方法。

（刘志国　屈　伸）

第六章　外源基因转移技术

外源基因与载体在体外重组形成重组 DNA 分子后，为了分析基因的表达、测定表达对细胞生长的影响、研究该基因的结构与功能、得到高表达的基因产物，通常需要将重组 DNA 分子通过特定的方法导入宿主细菌（细胞）；另外，由于研究的需要还可将包括基因组 DNA、寡核苷酸等外源 DNA 导入合适的受体细胞内。

宿主细胞可以分为原核细胞和真核细胞。原核细胞主要有大肠杆菌、链球菌、枯草杆菌、假单胞菌、放线菌等。真核细胞主要有酵母菌，近年来许多哺乳动物细胞（如猴肾细胞）及一些昆虫细胞也被作为宿主细胞使用。

质粒 DNA 或其他载体与目的基因 DNA 构建的重组载体导入宿主细胞的过程称为转化（transformation），而噬菌体、病毒或以它作为载体构建的重组载体导入宿主细胞的过程称为转染（transfection）。有时以噬菌体作为载体导入外源性 DNA 的过程也称为转导（transduction）。

习惯上，将微生物（如细菌、酵母）和植物细胞中进行的基因转移也称为转化，在哺乳类或其他动物细胞则称为转染（transfection）。本章着重介绍各种转化和转染方法的基本原理、实验技术和注意事项。

第一节　外源基因导入原核细胞的常用方法

一般来说，转化作用就是一种基因型细胞从周围介质中吸收来自另一种基因型细胞的 DNA，进而使原来细胞的遗传基因和遗传性发生相应变化的现象。在基因工程中，将重组 DNA 分子导入特定的受体细胞（宿主细胞），使之获得新的遗传性状（这种遗传性状的改变通常是持久性的，可分为遗传型和表型两方面）的一种手段称为重组 DNA 分子的转化。它是微生物遗传、分子遗传、基因工程等研究领域的基本实验技术。

转化现象是一个自然存在的过程。然而，在自然界发生的转化与 DNA 重组技术中使用的转化在原理和方式上均显著不同。在 DNA 重组技术中通过使用各种人工的方法将外源基因导入受体细胞，实现体外的转化。不同宿主细胞转化的方法各不相同。本节主要讨论细菌尤其是大肠杆菌的常用转化方法，包括物理和化学方法，如 $CaCl_2$ 转化方法、电穿孔转化方法等。在讨论这些方法之前，首先介绍自然存在的转化现象及其规律。

一、自然条件下的转化现象

最早发现的转化现象是在 1928 年英国的细菌学家 Griffith 在肺炎球菌的研究中最先观察到转化现象，并在 1944 年由美国的 Avery 在首次开展的肺炎球菌的转化试验中形成了完整的转化概念。Avery 将杀死的有荚膜的光滑型（S 型）肺炎球菌的核提取物与活的无荚膜

的粗糙型（R 型）肺炎球菌一起混合放置后，结果产生了有致病性的有荚膜肺炎球菌，称为 S-R 变异转化试验；该实验明确了外源 DNA 是细菌转化的根本原因。之后，将受体细菌吸收的供体细菌的游离 DNA 片段称为转化因子。

现已发现多种细胞具有转化能力，不同生理条件下的受体菌接受外源 DNA 的能力差别悬殊，通常将易于接受外源 DNA 片段而实现转化的特殊的生理状态称为感受态。处于容易吸收外源 DNA 状态的细菌叫感受态细菌。感受态细菌吸收转化因子的能力很强，通常比一般细菌高千倍以上。细菌的转化能力受其自身遗传特性、菌龄、生理培养条件等诸多因素的影响。

具有转化能力的 DNA 片段常常是双链 DNA 分子，单链 DNA 分子很难甚至根本不能转化受体菌。进入受体细胞的 DNA 分子通过复制、表达实现遗传信息的转移，使受体细胞出现新的遗传性状。将经过转化后的细胞在选择培养基中培养，即可筛选出转化子（transformant），即带有异源 DNA 分子的受体细胞。

目前对革兰阳性细菌的转化过程已较为清楚，而对革兰阴性细菌的转化机制尚有很多疑问。以下将主要介绍革兰阳性细菌转化的局部原生质体化假说。该假说认为细菌的转化机制在于在感受态过程中，细菌表面发生了局部的原生质体化。

① 细菌感受态的形成：当转化因子在接近细菌细胞时，细胞分泌一种小分子量的激活蛋白，又称为感受因子，该蛋白能诱导与感受态有关的特征性蛋白质（如细菌溶素）的合成，使细菌胞壁部分溶解，局部暴露出细胞膜上的 DNA 结合蛋白和核酸酶等，这些变化有利于细胞接纳外源转化因子、实现转化。

② 转化因子的吸收：转化因子（双链 DNA）与受体细胞膜上的 DNA 结合蛋白特异性结合，然后激活邻近的核酸酶。核酸酶降解双链 DNA 的一条链，而另一条链则被吸收到受体菌中。单链 DNA 或 RNA、双链 RNA 以及 DNA-RNA 杂合双链因不能与膜上特异蛋白质结合，而不能被细菌吸收。

③ 转化因子与宿主染色体 DNA 的整合：进入受体细胞的单链 DNA 与另一种游离的蛋白因子结合，形成整合复合物前体结构，它能有效地保护单链 DNA 免受各种胞内核酸酶的降解，并将其引导至受体菌染色体 DNA 处，通过同源重组，转化因子的单链 DNA 与宿主染色体 DNA 的同源区域置换，使转化因子插入到宿主 DNA 中，形成稳定的整合。

④ 转化子的形成：整合的转化因子随宿主染色体 DNA 一起复制，并通过细胞分裂而形成新的转化子。

革兰阴性细菌细胞因其表面的结构和组成与革兰阳性细菌不同，自然条件下的转化较为困难。目前了解到有些革兰阴性细菌（如嗜血杆菌和奈氏杆菌等）可以完整的双链形式吸收转化因子，被吸收的双链 DNA 与相应的 DNA 结合蛋白结合，可以不受核酸酶降解。另外一些革兰阴性细菌（如大肠杆菌）则难以吸收转化因子或细胞内存在降解线性转化因子的核酸酶系统。正因为如此，在 DNA 重组实验中通常采用人工的方法制备感受态细菌，直接将重组的 DNA 导入细胞内以及选择合适的宿主菌，以避免宿主菌对外来 DNA 的降解作用，确保较高的转化效率。

二、受体细胞的选择

选择受体细胞主要考虑细菌的转化效率和生物安全性。一般，野生型细菌对外源 DNA 的转化效率较低，并且有可能感染其他生物种群，因此不能作为基因工程的受体菌。目前使

用的受体菌必须是经过遗传改造的安全、高效的菌株，这种受体菌具有以下特点。

1. 限制修饰系统缺陷

野生型细菌一般具有针对外源 DNA 的限制和修饰系统，即能够降解外源 DNA，维持其自身遗传性状的稳定。这一特性对外源重组 DNA 转化受体菌极为不利，它会极大地降低转化效率。因此，在克隆操作中，为了提高外源 DNA 分子的转化效率，就需要选用限制系统缺陷型的受体细胞。对大肠杆菌来说，其限制系统主要由 *hsd*R^- 基因编码，因此应选择具有 *hsd*R^- 遗传表型的大肠杆菌菌株，能大大增加外源 DNA 的转化效率。

2. 重组系统缺陷

进入宿主细胞内的外源 DNA 分子能自发地与宿主染色体 DNA 发生同源重组。这一反应对于以 DNA 重组分子自主复制为基础的外源基因克隆、扩增以及表达来说极为不利，应尽可能避免。在大肠杆菌的同源重组过程中，*rec*A 基因编码的 RecA 重组蛋白能促进 DNA 分子之间的同源联会和 DNA 单链交换，为重组所必需。选择 *rec*A 缺陷（*rec*A^-）型大肠杆菌细胞可使重组频率比正常降低 10^6 倍。

3. 转化亲和型

用于基因工程的受体细胞必须对重组 DNA 分子具有较高的可转化性，通常利用遗传诱变技术改变受体细胞壁的通透性，从而提高转化效率。如果用噬菌体 DNA 作载体构建的 DNA 重组分子转染受体菌，则要求受体细胞膜上必须具有噬菌体的特异性吸附受体，如大肠杆菌膜蛋白 LamB 是 λ 噬菌体吸附所必需的结构等。

4. 遗传表型具有互补性

受体细胞必须具有与载体所携带的选择标记互补的遗传性状，方能使转化细胞的筛选成为可能。常用的遗传互补性状包括抗生素的抗性及营养缺陷表型。这样，当重组分子转入受体细胞后，载体上的标记基因赋予受体细胞特殊的表型特征，以区分转化细胞与非转化细胞。更为理想的受体细胞是具有与外源基因表达产物活性互补的遗传特征，这样便可直接筛选到外源基因表达的转化细胞。

5. 感染寄生缺陷型

相当多的细菌对其他生物（尤其是人和牲畜）具有感染和寄生效应，重组 DNA 分子导入这些细菌受体后，极有可能随着受体菌的感染寄生作用而广泛传播，如果外源基因对人体和牲畜有害，则会导致一场灾难，因此从生物安全的角度考虑，受体细胞不能具有感染寄生性，当然，利用基因工程手段制造生物武器例外。

受体细胞的上述特点对外源基因的克隆与表达来说都是十分重要的，它直接关系到基因克隆工作的成败。对于以改良生物物种为目的的克隆工作而言，受体细胞的选择则没有太大的选择余地，此时的受体细胞只能是待改良的生物物种本身。

三、转化方法

（一）$CaCl_2$ 诱导大肠杆菌转化法

目前常用的感受态细胞制备方法有 $CaCl_2$ 和 $RbCl_2$ 法，$RbCl_2$ 法制备的感受态细胞转化效率较高，但 $CaCl_2$ 法简便易行，且其转化效率完全可以满足一般实验的要求，制备出的感受态细胞暂时不用时，可加入占总体积 15% 的无菌甘油，于 −70℃ 保存（半年），因此 $CaCl_2$ 法使用更为广泛。

1970 年 Mandel 和 Higa 首先发现用氯化钙溶液处理细菌（K12 大肠杆菌），经短暂热休

克后容易被λ噬菌体 DNA 感染，随后 Cohn 于 1972 年进一步证明了质粒 DNA 用同样的方法也能进入细菌。随后，建立了用 $CaCl_2$ 化学法处理大肠杆菌，用人工制备的感受态菌，使外界介质中的 DNA 转入大肠杆菌的方法。该方法的转化效率可达 $10^5 \sim 10^6$ 个转化子/μg DNA。环状 DNA 比线状 DNA 分子转化效率要高 1000 倍左右。此法已广泛应用于常规的基因克隆的转化试验中，该法的关键在于必须使用对数生长期的细菌，并尽量保证在低温下操作。

将对数生长期的细菌在 0℃下，用预冷的 $CaCl_2$ 溶液低渗处理，使菌体的细胞壁和细胞膜通透性增加，菌体膨胀成球形。此时用于转化的 DNA 可形成抗 DNase 的羟基-钙磷酸复合物黏附于细菌表面，经短暂 42℃的热休克（热激反应）后，不仅使介质中的 DNA 易于进入细菌的细胞内，而且不易被菌体中的 DNase 降解。虽有报道热冲击对转化不起作用，但许多人仍习惯采用。为了提高转化效率，实验中要考虑以下几个重要因素。

① 选择合适的菌株。一般实验室选用经常使用而且认为转化效率高的菌株。为了防止外源 DNA 被菌体内限制性修饰系统的内切酶消化降解，通常选用宿主限制系统缺陷型（*hsd*R^-）或宿主修饰系统缺陷型（*hsd*M^-）的大肠杆菌菌株作为宿主。若用于高效表达目的基因，往往还需对不同宿主的表达效率通过比较而获得。

② 细胞生长状态和密度。受体菌从超低温（−70℃）或−20℃甘油保存状态下的菌种中取出，培养致敏，其转化率要比 4℃短期保存的菌株转化率高，因此最好不要用经过多次转接或储于 4℃的培养菌；但−70℃状态保存时间过长，也会降低转化率。

细胞生长密度以刚进入对数生长期时为好，可通过监测培养液的 OD_{600} 来控制，静止期的细胞不能被转化。DH5α 菌株的 OD_{600} 为 0.5 时，细胞密度在 5×10^7 个/ml 左右（不同的菌株情况有所不同），这时比较合适。密度过高或不足均会影响转化效率。

③ 质粒的质量和浓度。环状 DNA 要比线状 DNA 的转化率高，因而用于转化的质粒 DNA 应主要是超螺旋态 DNA（cccDNA）。小的环状 DNA 要比大的环状 DNA 转化率高。

转化效率与外源 DNA 的浓度在一定范围内成正比，但当加入的外源 DNA 的量过多或体积过大时，转化效率就会降低。1ng 的 cccDNA 即可使 50μl 的感受态细胞达到饱和。一般情况下，DNA 溶液的体积不应超过感受态细胞体积的 5%。

④ 试剂的质量。所用的试剂，如 $CaCl_2$ 等均需是最高纯度的（GR 或 AR），并用超纯水配制，最好分装保存于干燥的冷暗处。

在转化体系中应避免有锌离子的混入，因为锌离子可以强烈地抑制转化。

⑤ $CaCl_2$ 法 0℃放置时间的影响。研究发现细菌经 0℃ $CaCl_2$ 处理后转化率随时间的推移而增加，24h 达最高转化率，而后转化率逐渐下降。此法制备的感受态菌需新鲜制备，当天使用有效。通常不采用 4℃保存过久的感受态菌。

⑥ 防止杂菌和杂 DNA 的污染。整个操作过程均应在无菌条件下进行，所用器皿，如离心管、tip 头等最好是新的，并经高压灭菌处理。所有的试剂都要灭菌，且注意防止被其他试剂、DNA 酶或杂 DNA 污染，否则均会影响转化效率或杂 DNA 的转入，给以后的筛选、鉴定带来不必要的麻烦。

⑦ $MgCl_2$ 可以代替 $CaCl_2$ 进化转化。这种转化方法可以不经过细胞热休克过程，即一步操作可制备感受态细菌，但仅适合环状 DNA 转化，其应用范围较小。

⑧ 除了经典的 $CaCl_2$ 致敏转化法外，为了提高转化率或针对不同的受体菌株，有人联合锰离子（$MnCl_2$ 法）、铷离子（$RbCl_2$ 法）、钴离子（$CoCl_2$ 法）及二甲基亚砜（DMSO）

等，可以使转化率提高很多（100～1000 倍）。

目前，Ca^{2+} 诱导法已成功地用于大肠杆菌、葡萄球菌以及其他一些革兰阴性菌的转化。转化时常见问题及解决办法见表 6-1。

表 6-1　转化时常见问题及解决办法

转化率低的原因	解决办法
DNA 不纯	除去 DNA 溶液中的酚、蛋白质、去污剂和乙醇
DNA 量过多	每 100μl 感受态细胞中加入 DNA 的量为 1～10ng，DNA 样品体积不超过 10%
连接反应不完全	优化连接反应条件。为检测连接反应，可先进行预实验，转化经酶切后再连接的载体 DNA
连接反应物抑制转化	加感受态细胞前，将连接反应物用 1mmol/L Tris-Cl(pH 7.5) 和 1mmol/L EDTA 稀释 5 倍
抗性表达差	改用 SOC 为表达培养基。若 LB 为表达培养基，转化效率往往会下降 2～3 倍。在转化细胞涂布前。应至少培养 30min。使抗性表达；如用氨苄西林筛选。转化细胞必须稀释后再涂布，否则抗性菌落周围的一些非抗性菌落也会生长，用 50～100μg/ml 的氨苄西林可减轻这一问题
感受态细胞保存不当	感受态细胞存于－70℃约可保存 6 个月。不要将感受态细胞储存于液氮中。分装体积应适当小些，尽量减少冻融次数；冰浴时感受态细胞应于干冰-乙醇中快速冷冻。为获得最高的转化率，感受态细胞溶解后最好不再冰冻
操作感受态细胞不当	冰冻感受态细胞应在冰浴中溶解，一旦溶解后应立即使用，不要涡旋振荡
热激不当	根据预实验，固定菌株、管型号、反应体积及热激的最适温度和时间。如改变管子型号和反应体积，应重新调整温度和时间

（二）电穿孔驱动的完整细胞转化

当细胞置于非常高的电场中，细胞膜就变得具有通透性，能让外界分子扩散进细胞内，这一现象称为电穿孔（electroporation）。电穿孔技术最早用于将外源 DNA 导入真核细胞。1988 年 Dower 等将其用于大肠杆菌的转化。目前，运用这一技术，许多物质（包括 DNA、RNA、蛋白质、药物、抗体和荧光探针）都能转入细胞。作为一种基因转导方法，电穿孔已被广泛用于各种细胞类型，包括细菌、酵母、植物和动物细胞；而且，它还能作为注射方法（称为电注射），把各种外源物质引入活细胞。

电穿孔法的基本原理是在电场脉冲的作用下，受体细胞的细胞膜上形成一些微孔通道，使得 DNA 分子直接与裸露的细胞膜脂双层结构接触，促进 DNA 的吸收。电穿孔转化方法的具体操作程序因转化细胞的种属而异。转化效率受电场强度、脉冲频率、脉冲时间等因素的影响。仔细选择转化条件和参数，可达到每微克 DNA 转化获得 10^9～10^{10} 转化子的转化效率。对于大肠杆菌来说，大约 50μl 的细菌与 DNA 样品混合后，选用 25μF、2.5kV 和 200Ω 的电场强度处理 4.6ms，即可获得理想的转化效率。用电穿孔法制备感受态细胞要比用 $CaCl_2$ 转化方法容易，并可以转化较大的重组质粒（＞100kb），尽管对大质粒的转化效率只有小质粒（约 3kb）的千分之一，但比上述 $CaCl_2$ 方法和 PEG 转化方法理想，因为这两种方法几乎不能转化大于 100kb 的质粒 DNA。

与其他常用的导入外源物质的方法相比，电穿孔具有很多优点。首先，不必像显微注射那样使用玻璃针，不需要技术培训和昂贵的设备，可以一次对成百万的细胞进行注射。第二，与用化学物质相比，电穿孔几乎没有生物或化学副作用。第三，因为电穿孔是一种物理方法，较少依赖细胞类型，因而应用广泛。实际上，对大多数细胞类型，用电穿孔法基因的转移效率比化学方法高得多。

（三）PEG 介导的细菌原生质体转化

某些微生物，如酵母菌、霉菌和棒状杆菌，很难得到感受态细胞，需要首先制备原生质

体来增加细胞对 DNA 的吸附能力，然后再进行转化。另外一些微生物如枯草杆菌，既能用感受态细胞转化，又可以用原生质体转化。原生质体转化方法不仅适用于芽孢杆菌和链霉菌等革兰阳性细菌，而且对植物等真核细胞也有效，只是不同种属的生物细胞，其原生质体的制备与再生的方法不同。

基本的操作步骤：在高渗培养基中生长至对数生长期的细菌，用含有适量溶菌酶的等渗缓冲液处理，剥除其细胞壁，形成原生质体，这样形成的原生质体丧失了一部分定位在膜上的 DNase，有利于双链环状 DNA 分子的吸收。此时，加入含融合剂聚乙二醇（PEG）的等渗溶液和待转化的 DNA 样品，混匀后离心除去聚乙二醇，将菌体涂布在固体再生培养基上，经培养后再生出细胞壁，获得转化细胞。

原生质体转化法的原理是：先除去细胞壁，待外源 DNA 进入细胞后放在有利于再生细胞壁的培养基中培养。原生质体转化不需要感受态细胞，而且非完整的质粒也能进行转化，如在克隆时留有缺口的质粒以及非超螺旋环状分子都能进行转化，且转化效率尚可。原生质体的制备及再生过程繁琐，而且经常融合产生二倍体、三倍体。

（四）通过接合作用传递质粒 DNA

F 质粒（F 因子）又称性质粒，除了含有自我复制基因外，还带有一套接合转移的基因。凡携有 F 质粒的细菌称 F^+ 菌株，不带有 F 质粒的细菌称 F^- 菌株。F^+ 菌株（供体菌）一旦与 F^- 菌株（受体菌）发生接触，F^+ 菌株可通过性菌毛将 F 质粒转给 F^- 菌株，使 F^- 菌株转变成 F^+ 菌株。质粒由 F^+ 向 F^- 菌的转移过程叫接合转移，又称质粒的接合传递。凡带有在染色体上整合 F 质粒的细菌又称高频重组株系（Hfr 株系）。Hfr 株系不稳定，F 质粒可发生接合转移。F^+ 菌株可失去 F 质粒，F^- 菌也可获得 F 质粒，其转移难以人为控制，通常不用作克隆载体。

（五）λ 噬菌体的转导和转染

噬菌体 DNA（或重组子）也可被感受态细胞摄取并产生有活性的病毒颗粒，这个过程被称为转染（transfection），transfection 字面上由 transformation 和 infection 组合而成，其特点是提纯的噬菌体 DNA 以转化的（而非病毒感染）途径进入细胞并表达后产生完整的病毒颗粒。

噬菌体为细菌的病毒，以噬菌体的病毒颗粒为媒介，将外源 DNA 导入细菌的过程称转导。

噬菌体转染受体细胞与上述质粒转化大肠杆菌的被动吸收方式不同，它是一种主动的感染方式，因此其转化效率很高。每微克野生型 λ 噬菌体在良好的包装条件下转染大肠杆菌细胞，可形成 10^8～10^9 个 pfu（plaque forming unit，噬菌斑形成单位），重组的 λ 噬菌体转化效率有所降低，但也可达到 10^6～10^7 个 pfu。另外，λ 噬菌体装载的外源 DNA 分子比质粒装载的大，这些特点使得 λ 噬菌体转染大肠杆菌的转化方式被广泛用于基因组文库和 cDNA 文库的构建。

与质粒不同的是，λ 噬菌体转染大肠杆菌前需要进行体外包装，使之成为具有感染活力的噬菌体颗粒。用于体外包装的蛋白质可以直接从大肠杆菌中制备，现已商品化。包装所需的蛋白成分通常分为功能互补的两部分，独立放置。包装时，将两种包装组分与重组 λ DNA 分子混合后，即可有效地包装，缺少其中任何一种蛋白组分均不能包装成有感染活力的噬菌体颗粒。包装操作过程相当简单：将 λ DNA 与外源 DNA 片段的连接反应液，与两种包装蛋白组分混合，在室温下放置 1h，加入 1 滴氯仿，离心除去细菌碎片，即得重组噬

菌体颗粒的悬浮液。适当稀释后与处于对数生长期的大肠杆菌受体细胞混合涂布，过夜培养即可用于筛选与鉴定。

四、转化率及其影响因素

1. 转化率的定义

转化率是转化效率的评估指标。通常有两种表示方式。一是在待转化 DNA 分子数大于受体细胞数的条件下，转化细胞与细胞总数之比，例如在标准条件下，利用 Ca^{2+} 诱导法转化质粒 DNA 的最大转化率为 10^{-3}，即平均每 1000 个受体细胞中有一个细胞接纳了质粒 DNA，如果假定处于感受态的受体细胞能 100%地接纳 DNA 分子，则这种转化率直接反映了受体细胞中感受态细胞的含量。

转化率的另一种表示方式是在受体细胞数大大超过质粒 DNA 分子数时，每微克 DNA 转化受体细胞所产生的克隆数。由于在一般的转化实验中，每个受体细胞只能接纳一个 DNA 分子，因此上述转化率可换算为每微克 DNA 进入受体细胞的分子数。例如，pUC18 对大肠杆菌的转化率为 $10^8/\mu g$ DNA，其含义是每微克 pUC18 只有 10^8 个分子能进入受体细胞，由于 1μg pUC18 包含 3.4×10^{11} 个分子 [$6.02\times10^{17}/(2686\times660)$]，因此，每 3.4×10^{11} 个分子中有 10^8 个 pUC18 分子（即 3400 个 pUC18 分子中只有一个分子）进入受体细胞。

转化后在含抗生素的平板上长出的菌落即为转化子，根据平皿中的菌落数可计算出转化子总数和转化频率，公式如下。

转化子总数=(菌落数×稀释倍数×转化反应原液总体积)/涂板菌液体积

转化频率(转化子数/μg 质粒 DNA)=转化子总数/质粒 DNA 加入量(μg)

感受态细胞总数=(阳性对照组菌落数×稀释倍数×菌液总体积)/涂板菌液体积

感受态细胞转化效率=转化子总数/感受态细胞总数

2. 转化率的用途

通过已知的转化率参数可以帮助评估 DNA 重组实验的规模。例如，用 λ 噬菌体构建文库，λ DNA 对大肠杆菌的感染率（即转化率）为 $10^7/\mu g$ DNA，假定重组系统的连接效率为 20%，产生的重组 λ DNA 的转化效率为未重组的 λ DNA 的 1/100 以下，则欲获得 10^4 个含有重组 DNA 分子的克隆，至少应投入 0.5μg 的载体 DNA，其计算方法如下：$10^4/(20\%\times10^{-2}\times10^7)$。按外源 DNA 片段与载体 DNA 分子数 10 比 1 的要求，即可推算出外源 DNA 片段的用量。如果转化培养液全部涂板筛选，理论上可获得 5×10^4 个转化克隆（或噬菌斑），若使每块平板上含有 1000 个克隆，则需涂布 50 块平板。

3. 转化率的影响因素

转化率的高低意味着转化所获得的转化子数目的多少，很显然，在同等条件下，转化率越高，转化过程中获得的转化子数也就越多，反之，转化率低，意味着转化过程获得的转化子数少。为了获得高转化率应从重组 DNA、受体细胞和转化手段三个方面考虑。

(1) 重组 DNA 分子与转化效率　首先，不同的重组 DNA 分子对同一受体菌细胞进行转化的效率是不一样的。一般规律是：第一，小分子量的质粒转化率较高，大质粒通常转化率较低，超过 30kb 的重组质粒则很难进行转化；第二，质粒与受体细胞的亲和性也可影响转化，亲和性强的质粒转化效率高于亲和性弱的质粒；第三，结构致密的天然超螺旋结构的质粒转化效率高于体外构建的重组质粒，双螺旋闭环结构质粒转化效率高于结构松散的开环

结构或线性结构的质粒。第四，重组 DNA 分子的浓度和纯度对转化率也有影响，在 10ng/100μl 以下的 DNA 浓度范围内，转化效率与 DNA 分子数成正比关系。

（2）受体细胞与转化效率　同一重组质粒 DNA 分子转化不同的受体菌细胞，转化率不同。例如，每 μg 完整的 pBR322 DNA 分子转化大肠杆菌 JM83 时，转化率不高于 10^3/μg DNA，但若转化 ED8767 菌株，则可获得 10^6/μg DNA 的转化率，若转化大肠杆菌 X1776 菌株，转化率更是高达 10^8/μg DNA。因此，受体细胞的选择对于重组 DNA 分子的转化也是极为重要的一环。

（3）转化方式对转化的影响　选择不同的转化方法，转化率也有所差异（表 6-2）。在上述三种转化方法中，以电穿孔转化方法的转化率最高，达 10^8～10^9；$CaCl_2$ 诱导转化法约为 10^6～10^7；而接合转移法最低，仅为 10^4～10^5，但它适于前两种方法难以转化的受体菌。在转化操作方面，受体细胞的预处理或感受态细胞的制备对转化率影响最大。在 $CaCl_2$ 转化法中，制备感受态所用大肠杆菌的菌龄、$CaCl_2$ 处理时间、感受态细胞的保存期以及热处理时间均是重要的影响因素。在电穿孔转化法中，各种参数以及对受体细胞的预处理均可影响转化率，仅冰冻甘油预处理对转化率的影响，就比不经此预处理的转化率高 10～100 倍。

表 6-2　大肠杆菌常用转化方法的比较

方　法	转化频率/（转化子/μg DNA）	方　法	转化频率/（转化子/μg DNA）
$CaCl_2$ 方法	10^6～10^7	电穿孔法转化	10^8～10^9
原生质体转化	10^5～10^6	接合转化	10^4～10^5
λ噬菌体转化	10^7～10^8		

第二节　外源基因导入真核细胞的方法

由于真核生物的细胞结构、基因组成和基因表达比较复杂，适用于原核细胞的转基因转移方法大多难以有效地应用于真核生物。随着分子生物学技术的发展，一系列适用于真核生物的转基因方法已被建立并逐步完善起来。目前，基因转移技术已广泛用于基因的结构和功能分析、基因表达与调控、基因治疗与转基因动物等研究。

将某一特定的靶基因传递到靶细胞并在细胞内表达，需要应用某一基因运载系统，目前将这一系统概括为非病毒类和病毒类两大类，其区别在于将目的基因转入靶细胞并指导其进入细胞核的方法不同。病毒载体法是以病毒作为外源 DNA 导入的工具，通过病毒感染将外源 DNA 导入细胞。这种方法的优点在于它具有较高的转染效率，在一定条件下能实现对特定细胞的靶向性。该方法目前主要用于基因治疗的研究中。本节将介绍目前基因转移的非病毒方法；基因转移的病毒方法（基因转染技术）将单独在下一节中做重点介绍。

基因转移的非病毒方法又可分为物理和化学的方法以及受体介导的基因转移的方法。物理的方法包括直接注射法、显微注射法、电穿孔法、微粒子轰击法（基因枪），化学的方法包括磷酸钙共沉淀法、DEAE-葡聚糖和脂质体转染法等，受体介导的基因转移是近几年来发展起来的一种新的基因转移技术。

根据不同的实验目的，转染实验又可分为瞬时转染和稳定转染。除 DEAE-葡聚糖介导

法只适合瞬时转染外，其他各种物理化学方法既可用于瞬时转染又可用于稳定转染；而不同的病毒载体法，有的适用于稳定转染（如逆转录病毒载体法），有的适用于瞬时转染（如腺病毒载体法）。为了获得稳定转染的细胞，须使转染用的质粒载体 DNA 中含有合适的耐药性基因（如新霉素抗性基因 *neo*）表达元件，并在转染后对转染细胞进行相应的药物筛选（如对于 *neo* 基因常用 G418 筛选。G418 是一种与新霉素、卡那霉素相似的氨基糖苷类抗生素）。

对于不同的受体细胞，转染的方法也不尽相同。对于贴壁细胞来说，磷酸钙及脂质体介导的基因转染方法为首选；病毒载体法也较常用，该法转染效率高，但存在生物安全性问题，对实验室的防护要求较高。而对于非贴壁细胞来说，可用电穿孔法或脂质体介导的基因转移方法进行转染。

综上所述，转染方法有多种，应根据不同的细胞，贴壁或悬浮细胞等选用不同的方法，其目的是要达到需要的转染效率。影响转染效率的因素有多种，包括转染方法、操作技术、质粒 DNA 的纯度、靶细胞的生长状态等。本节着重介绍各种转染方法的基本原理和注意事项。

一、物理方法

1. 电穿孔法

其原理在原核细胞的基因转移中已提及，在此仅做补充说明。

细胞膜的作用就像是一种电容器，除非使用离子通道，否则电流是无法通过的。但在高压电场的作用下，细胞膜因发生临时性破裂所形成的微孔，足以使大分子以及像 ATP 这样的小分子从外界进入细胞内部，或是反向流出细胞。细胞膜上微孔的关闭是一种天然衰减的过程，在 0℃下，这种过程会被延缓进行。

在微孔开启期间，细胞外环境中的核酸分子便会穿孔而入，并最终进到细胞核内部。具有游离末端的线性 DNA 分子，易于发生重组，因而更容易整合到寄主染色体，形成永久性的转化子。超螺旋的 DNA 比较容易被包装进染色质，因此一般说来它对于瞬时基因表达的实验更为有效。

电穿孔法与其他常用的基因转移方法相比，具有许多优点，如较少依赖于细胞类型、应用广泛（包括微生物细胞的基因导入都能用电穿孔的方法）。电穿孔既可用于瞬时表达也可用于稳定转化，转染效率比化学方法高得多（电穿孔法可使大多数细胞系和一些原代细胞的瞬时转染率达到 90%，而得到稳定转染细胞的频率为 $10^{-3}\sim10^{-4}$）。与化学方法或病毒载体法进行比较，电穿孔法几乎没有生物或化学副作用等。

影响电穿孔法基因转移效率的因素有以下几种。

① 电脉冲的最大电压及其持续的时间：电压太小或持续时间太短，转染效率低；电压太大或持续时间太长，电击后细胞不能存活。

与电脉冲的形状关系不大（电穿孔仪可产生三种电脉冲，即方波脉冲、指数衰减脉冲和振荡电场脉冲。其中，以指数衰减脉冲又称电容放电脉冲的发生器结构简单、价格便宜，最为常用）。

实验表明，不同种类的细胞需要不同大小的电击电压，才能得到有效转染，而且该电压值只存在于一狭窄范围内。

② 受体细胞的类型、转染用的 DNA 构象等。电穿孔法既适用于贴壁生长细胞，也适用

于悬浮生长细胞。电穿孔转染效率与 DNA 浓度则无多大关系。一般说来，每 10^7 个细胞所用 DNA 总量在 10～40μg 之间，可以得到很好的转染结果，而且 DNA 浓度同被细胞捕获的 DNA 数量之间存在着正比的线性关系。

DNA 构象对电穿孔转染效率的影响较大，一般认为线状 DNA 转染的成功率较环状 DNA 高，但对于瞬时转染也可采用环状 DNA。

③ 电击条件选择：不同细胞对电脉冲的敏感性不同，对不同类型的受体细胞，为获得最佳转染效率和细胞存活率，其所需的电脉冲调节参数如电压、电容、电脉冲持续时间等，都要进行预试验以确定其最佳的电击条件。目前，对动物细胞进行电穿孔，建议采用长脉冲/低电场方法；对于微生物细胞，则建议使用短脉冲/高电场方法。一般选择造成 20%～60%靶细胞死亡的脉冲调节参数为特定细胞的电击条件。

④ 电击缓冲液：细胞受电击后产生的孔洞使细胞质与电击缓冲液直接接触，电击缓冲液组分能直接进入细胞内，而细胞对缓冲液的离子组成和渗透压非常敏感，因此组分与细胞质类似的缓冲液有利于细胞电击后的存活，常用 Cytomix 缓冲液。与其他的电击缓冲液 PBS 和完全培养基相比，Cytomix 缓冲液最大程度地模拟了真核细胞内的离子环境；而且所含的 ATP 和谷氨酰胺能阻止细胞质组分渗漏，保护细胞膜抵抗氧化作用，有利于孔洞的重新闭合；此外，还可加大电击强度来获得最佳瞬时转染效率。

2. 显微注射法

通常根据 DNA 注射的不同方式，可以把显微注射区分为真正的显微注射（true micro-injection）法和“穿刺”（pricking）导入法两种。显微注射时，DNA 是由注射针直接注入细胞；而在穿刺中，DNA 是处在细胞周围培养基中，它通过穿刺形成的小孔进入细胞的内部，或是随穿刺的针头一道进入。

（1）显微注射法　在显微镜下，应用玻璃显微注射器，可以把重组 DNA 直接注射到哺乳动物细胞的细胞质或细胞核。该法适合于各种培养生长的细胞，但需要一定的设备和操作技巧。

用显微注射法转移基因的效率明显地超出其他方法。如把 tk DNA 注射到小鼠 tk^- 细胞，结果有 50%～100%的受注射细胞能瞬时表达胸苷激酶活性。获得稳定转化子的数量，取决于注射的 DNA 的性质。如一个对比实验，一组用 pBR322/HSV-1 tk 重组 DNA 注射，结果在 500～1000 个受注射细胞中，只有一个转变成稳定的转化子；另一组用连接着特定的 SV40 序列的 pBR322/HSV-tk 重组 DNA 注射，结果平均每 5 个受注射细胞中便有一个稳定的转化子，转化频率达 20%。

显微注射用的受体细胞，多数是沿培养皿贴壁生长的单层细胞。据试验结果推算，每个细胞注射的 DNA 样品平均体积只有 $(1\sim2)\times10^{-11}$/ml。这个数字只相当于 1 个细胞体积的 1%～2%。一个熟练的技术员，每小时可以注射 500～1000 个细胞，其中成功地导入外源 DNA 的约占 50%～100%。通过对 DNA 样品的适当稀释，便可以控制注射的 DNA 分子数。

（2）穿刺法　穿刺法是使用显微注射针穿刺细胞及细胞核，而把外源 DNA 导入培养细胞或细胞核的一种特殊的显微注射法。简单地说，这种方法的操作过程是取对数生长期的细胞，加入蛋白酶及 EDTA 溶液，于 37℃ 消化 15min；移去消化液，并用培养基漂洗细胞之后，在每个直径为 6cm 的培养皿中接种 $(1\sim5)\times10^3$ 个细胞，置 37℃继续培养过夜；在穿刺前弃去培养基，用 pH7.2 的磷酸缓冲液（PBS）漂洗 2 次，然后加入浓度为 5～1000μg/ml

的 DNA-PBS 溶液，覆盖于培养细胞表面；划出一个区域并做上标记，将该区内所有的细胞都进行穿刺，另有一部分细胞不穿刺供作对照；穿刺后，马上移去 PBS-DNA 溶液，换上正常培养基，随后再换上选择培养基。

3. 微粒子轰击法（基因枪）

微粒轰击又被称为基因枪（gene gun）技术、生物发射技术或高速微粒子发射技术，是近几年发展较快的转移方法。它使用高能微粒子轰击将 DNA 导入培养的细胞或活的哺乳动物组织内。亚微粒的钨和金能自发地吸附 DNA，将这些微弹以很高的速度轰击细胞。实验结果表明，用这种方法靶基因可在皮肤、肌肉、肝、胰、胃和乳腺等细胞中表达。

基因枪技术适用于植物、动物细胞的基因转染。其不足也在于需要专门仪器。

4. 光穿孔法

光穿孔法（optoporation）是利用激光产生的热量，使细胞壁产生孔洞，将外源物质导入细胞。Palumbo 等（1996）将一束蓝色氩激光通过 100 倍物镜作用于培养基中的细胞，在酚红存在下由于温度的变化，受照部位细胞壁通透性增加，从而使悬浮于培养基中的 DNA 进入细胞。受照部位直径可由照射时间和照射强度控制，通透性持续时间很短，1～2min 内即自动消失，细胞无明显伤害。

光穿孔优点是：①利用培养基中常规成分酚红作用，无需任何添加剂；②可以有选择地对细胞进行转染，即当不同的细胞混在一起时，只要形态上可区分，就可被光穿孔法选择性地转染。不足之处为该法需要专门仪器。

5. 冲击波法

冲击波法（shock wave permeabilization）是利用细胞受冲击后细胞膜通透性瞬时增加，使外源基因导入细胞。产生于体外的冲击波，还可使组织中多种接触 DNA 的细胞得到转染。这种 DNA 转移方法操作非常简单，被转移的 DNA 大小、序列可灵活多样，可用于细胞及活体局部的基因转移，非常安全，能应用于临床。

6. 裸 DNA 直接注射法

1990 年 Wolff 等首次注意到给小鼠肌内直接注射纯化的 DNA 或 RNA 重组表达载体，可使载体上的基因在局部肌细胞内表达，这种表达可持续数日或更长时间，且没有检出注射的外源核酸与宿主染色体整合，随后的大量动物实验都说明在合适的条件下，DNA 接种后可诱导机体产生细胞及体液免疫。于是，作为一种新的基因治疗手段——核酸免疫技术应运而生。

实验表明，横纹肌细胞（如心肌、骨骼肌）和胸腺滤泡细胞摄取 DNA 的能力较强，但实际上质粒 DNA 注射入肌组织后，最多只有 1%～2%的肌纤维被转染，DNA 的转染效率不仅与质粒 DNA 的大小、构型有关，还与肌细胞的状态有十分密切的关系。人们为提高肌细胞摄取 DNA 的能力采取了一些措施，如在质粒接种前先给予高渗蔗糖溶液、心肌毒素或麻醉剂（如利多卡因）的预处理。

二、化学方法

1. 磷酸钙共沉淀法

磷酸钙沉淀法是由 Graham 等于 1973 年首先建立的，最早用于将病毒 DNA 导入培养的单层细胞，以后逐渐成为基因转染最常用的方法之一。磷酸钙共沉淀法是将溶解的 DNA 加在 Na_2HPO_4 溶液内，再逐渐加入 $CaCl_2$ 溶液；当 Na_2HPO_4 和 $CaCl_2$ 形成磷酸钙沉淀时，

DNA 被包裹在沉淀中，形成 DNA-磷酸钙共沉淀物。当该沉淀物与细胞表面接触时，细胞可能是通过非特异性内吞作用将 DNA 导入其中。若要得到稳定的转化细胞（即外源基因整合到宿主细胞基因组），可将细胞在选择培养基中传代培养，一般需要 2 周左右才能得到稳定的转化细胞。

该法的优点是方法简单，而且可以进行共转化，即将不含选择标记的 DNA 和含选择标记的 DNA 放在一起形成混合的共沉淀物，一起导入细胞。

用该法转染的 DNA 既可是基因组 DNA，也可是质粒 DNA。当用该法转染质粒 DNA 时，可加入分离自特定细胞或组织的高分子量基因组 DNA（称为 carrier DNA），以提高转染的效率。

该法的不足是不太适于悬浮细胞的转染。但近来也已有成功的报道。最初磷酸钙共沉淀法的转染效率很低（10^{-6}～10^{-4}），后来经过多次改造，其稳定转染效率已超过 10^{-2}。

根据所使用的缓冲溶液系统不同，目前常用的磷酸钙沉淀法实验方案有两种：利用 HEPES 缓冲溶液系统和利用 BES 缓冲溶液系统。使用 HEPES 缓冲液可形成一种能直接沉积在细胞上的磷酸钙沉淀；而使用 BES 缓冲液可使沉淀在培养基中逐步形成并沉积在细胞表面。这种方法对于环状质粒 DNA 进行稳定转化尤其有效。对于转化线状 DNA、基因组 DNA 或进行瞬时表达，这两种方法效率无明显差异。两种方法均需要高质量的 DNA。

影响转染效率的主要因素如下。

① DNA 的纯度和构象。转染用的 DNA 需用 $CsCl_2$ 密度梯度离心或其他方法进行纯化，以除去 RNA、蛋白质等杂质，获得高纯度的超螺旋质粒 DNA；不纯的 DNA 会影响转染效率。一般认为，超螺旋状态的质粒 DNA 转染效果好。对于线状 DNA 来说，可能是由于磷酸钙-DNA 共沉淀物形成太慢，延迟了保证 DNA 免受核酸酶降解作用的时间，导致了其转染效率极低。

② DNA-磷酸钙共沉淀物中 DNA 的数量。一般说来，在磷酸钙转染实验中要使用高浓度的 DNA（10～50μg/ml）。在共沉淀物中，DNA 总浓度的变动会显著地影响细胞捕获 DNA 的效率。据推测，这很可能是因为 DNA 总量影响了磷酸钙沉淀物的性质，从而改变了可被细胞捕获的外源 DNA 的比例。

③ 沉积反应和形成的颗粒大小。DNA-磷酸钙比例不当，则沉淀过大，直接影响转染效率。

④ 共沉淀物与细胞接触的保温时间。DNA 磷酸钙共沉淀物同细胞接触的最适保温时间，因细胞的类型而异。某些类型的细胞，如 HeLa 细胞，保温 16h 后就能被有效地转染；而另外一些细胞在这样长的保温时间内，仍不能被有效转染。

⑤ 细胞的密度。在开始转染时细胞数量必须小于 10^6 个/培养皿（直径 10cm），以留出足够细胞扩增 2 倍以上的生长面积。

⑥ 培养的条件。3%CO_2 浓度是较关键的步骤，必须认真调适。

⑦ 甘油/二甲基亚砜（DMSO）处理可提高某些细胞（如 CHO 细胞）的转染效率，但二者对细胞有毒，时间应严格把握。

已发现有许多种化学“促进因子”，会影响转染 DNA 在哺乳动物受体细胞中的表达效率。促进因子通常是在 DNA-磷酸钙共沉淀物被细胞吸收 4～8h 之后再加入。有些促进因子如 DMSO 可十分明显地提高外源 DNA 在幼年仓鼠肾细胞（baby hamster kidney cell，BHK）和小鼠白血病细胞中的表达水平。实验中使用的促进因子的最有效浓度、最适的持

续作用时间以及在细胞转染之后加入的最佳时期，对于每种促进因子和每种细胞系而言，都必须重新测定。

2. DEAE-葡聚糖转染技术

二乙氨乙基葡聚糖（diethyl aminoethyl dextran，DEAE-dextran）简称 DEAE-葡聚糖，是一种在生理 pH 条件下带正电荷的高分子量的多聚阳离子试剂，能促进哺乳动物细胞捕获外源的 DNA，实现瞬时的有效表达。最初被用于辅助病毒感染靶细胞，后来发现它也能促进哺乳动物细胞摄入外源 DNA，因此被广泛用于基因转染。

DEAE-葡聚糖促进细胞捕获 DNA 的具体机理还不清楚，曾经提出过几种较为合理的解释。一种认为 DEAE-葡聚糖同 DNA 结合成复合物，可以保护 DNA 免受核酸酶的降解作用；另一种则认为 DEAE-葡聚糖可以同细胞膜发生作用，从而使 DNA 能够容易地穿过细胞表面进入细胞内部。

DEAE-葡聚糖与磷酸钙共沉淀法比较有以下几点不同。

① 它一般用于克隆基因的瞬时表达，不易形成稳定转化细胞系，因而不适用于细胞的稳定转染；

② 由于它对细胞有毒性作用，所以某些细胞系（如 BSC-1、CV-1、COS 等）用该转染技术时转染效率很高，而其他类型的细胞则效率不满意；

③ DEAE-葡聚糖法所需 DNA 量要比磷酸钙沉淀法少，且从不使用载体 DNA；

④ 用该法导入细胞的 DNA 突变率高于用磷酸钙沉淀法。

DEAE-葡聚糖转染技术具有方法简单、重复性高、特别适用于研究基因瞬时表达（transient expression）实验、转染效率高于磷酸钙法等优点。例如，用 SV40 DNA/DEAE-葡聚糖转染猿猴细胞，感染率可高达 25%，而 SV40 DNA-磷酸钙法则最多只能达 15%。这个方法与磷酸钙法不同，它不能产生稳定的转化细胞系。造成这种差别的原因可能是同 DNA 组装成微染色体有关。因为有实验报道，一旦质粒 DNA 进入经过 DEAE-葡聚糖处理的细胞，它就会迅速地组装成含核小体的微染色体，而对磷酸钙处理的细胞，质粒 DNA 则是聚积起来或是以沉淀形式存在。

所谓基因的瞬时表达，是指在 DNA 进入细胞后迅速发生的反应，因此可作为一种有用的基因分析系统。一般认为，像这类的基因瞬时表达实验，它之所以能够成功地用于研究基因的转录，原因在于高水平的瞬时表达能合成出足够数量的基因产物，供分析研究使用。待研究的目的基因很可能并不具备强启动子，所以使尽可能多的细胞都表达目的基因，以使基因的拷贝数尽可能地丰富是十分重要的。而 DEAE-葡聚糖转染的细胞往往可以获得高水平的表达，故此法很有用。

DEAE-葡聚糖转染主要采用两种不同的方式。一种是先将 DNA 直接同 DEAE-葡聚糖混合，形成 DNA/DEAE-葡聚糖复合物后，再用来处理细胞；另一种是受体细胞先用 DEAE-葡聚糖溶液预处理，然后再用来与转染的 DNA 接触。在这两种转染方式中，受体细胞在用 DNA 或 DNA/DEAE-葡聚糖复合物处理前，都要先用等渗溶液漂洗，除去培养液中的血清成分。

依据受体细胞的不同特性和所采用的实验方式，加入的 DEAE-葡聚糖溶液的浓度范围变动在 100～1000μg/ml。一般在低浓度（200μg/ml）和较长的持续处理时间（8h 或 16h）的条件下，哺乳动物细胞捕获外源 DNA 的效率比较高。若在 DEAE-葡聚糖处理之后，再加入 DMSO 等促进因子，还可以进一步提高效率。

影响DEAE-葡聚糖转染效率的因素有许多，其中以细胞的数量、DNA的浓度和DEAE-葡聚糖的浓度最为重要。大体说来，按每个直径10cm的培养皿中加入1～10μg的转染DNA，并使用每毫升培养基含100～400μg DEAE-葡聚糖的溶液，对于绝大多数类型的细胞，都能得到很好的转染效率。

由于DEAE-葡聚糖对有些细胞具有毒性，因此DEAE-葡聚糖的浓度和处理细胞的时间尤为重要。据此，DEAE-葡聚糖进行转染实验的方案也有两种：用高浓度（0.5～1mg/ml）短时间（0.5～1.5h）；或用低浓度（100～400μg/ml）长时间（4～8h）两种方法来进行转染。随细胞的种类不同，所用DEAE-葡聚糖的最适浓度和最佳处理时间也不同，具体的时间可通过观察细胞与DNA/DEAE-葡聚糖接触后形态的改变来确定，若细胞依然贴壁生长，则继续培养；若细胞形态与加入DEAE-葡聚糖前不同，甚至不再贴壁，应终止培养。由于DEAE-葡聚糖对细胞有较强的毒性，所以采用低浓度长时间处理的方法比较可靠。

一般认为，同时加入氯奎（chloroquine）有保护DNA/DEAE-葡聚糖不被溶酶体降解的作用，可以进一步提高转染效率。用它作载体可以将外源DNA导入培养的哺乳动物细胞。

3. 脂质体法

脂质体（liposome）是一种人造膜，是用人工方法将磷脂在水溶液中形成一种脂质双层包围水溶液的脂质微球，外周是脂双层，内部是水腔。脂质分子在水溶液中形成脂质微球时，可将外源DNA包入脂质微球中，因此它可作为一种运载工具，通过脂质体膜与体细胞的相互作用（包括膜融合、吞噬等），把含有特殊功能的外源DNA导入细胞中。通过这种很简单的方法，外源DNA在多种不同的细胞中获得了有效表达。这种由脂质体介导的使外源DNA导入细胞的方法，叫做脂质体载体法，又叫做脂质转染法。

脂质体载体法的转染效率，比裸露DNA高出100倍；虽然细胞在加入脂质体之后，用高浓度的聚乙二醇（PEG）或甘油处理，可使脂质体DNA感染性提高10～200倍，而与DEAE-葡聚糖法相比至少还要低10倍。

尽管脂质体载体法的转染效率相对地低于其他方法，但这并不妨碍它在哺乳动物基因转移中的实用价值，因为它也具有许多优点。例如，制备程序比较简单，可以通过多聚碳酸酯滤膜（polycarbonate filter）消毒灭菌，用它包装的DNA在4℃可长期保存不失活性，以及毒性低、包装容量大、可保护DNA免受核酸酶的降解作用，以及结果可靠和可重复性强的特点。脂质体能在体内或体外提供运载外源性遗传物质进入细胞的载体，脂质体介导的基因转移的最大优势在于能在活体内应用。

以脂质体为载体的基因转移过程通常是先用PEG处理培养的动物细胞，使之易于吸收周围培养基中的脂质体。在正常情况下，每个细胞平均可吸收1000个左右，若加甘油或DMSO等促进因子，吸收量还要增多。

脂质体法如同显微注射法和电穿孔法一样，也可以用来将外源的蛋白质导入受体细胞，为在活细胞内研究蛋白质的功能提供有用的途径。运用这样的技术，不仅可以是将纯化的蛋白质直接导入活细胞，以分析其在天然环境条件下的功能效应，而且还可以将蛋白质的抗体导入细胞，阻断相应蛋白质（抗原）的功能。

市场上已有许多脂质体转染试剂出售。制备用于基因转染的脂质体时常用的脂质有：胆固醇、磷脂酰胆碱、磷脂酰甘油、磷脂酰丝氨酸、二油酰磷脂酰乙醇脂、磷脂酰乙醇脂、二油酰磷脂酰胆碱、油酸等。近年来新材料更是层出不穷。脂质体的脂双层分子可以掺入糖脂和抗体等归巢装置（homing devices），提高脂质体的靶向性。这些新材料和新方法为高效转

染哺乳动物细胞提供了广泛的选择基础。

单阳离子脂质（如 DOTMA、DOTAP 等）及中性磷脂（常见的为 DOPE）可与 DNA 形成类脂复合物，从而使 DNA 更有效地进入细胞。上述过程形成的类脂-DNA 复合物，并非通过双层脂质包裹形成真正的脂质体，而是阳离子脂质与 DNA 通过离子间相互作用，形成由阳离子脂质包裹 DNA 的颗粒。类脂-DNA 复合物的大小、所带电荷及转移的效率还可通过改变复合物中类脂的组成（如采用 DOSPA、DOSPER 等多聚阳离子脂质）及其与 DNA 的比例进行优化。研究表明，在体外，多聚阳离子脂质包裹的 300～700nm DNA 颗粒比单阳离子脂质包裹的 50～100nm DNA 颗粒的基因转移效率高得多。目前，阳离子脂质转染试剂已商品化，其转染效率比其他转染方法更高（可达磷酸钙沉淀法的 5～100 倍），重复性更好。

脂质体是唯一被美国 FAD 批准进行临床试验的基因转移的物理化学方法。一般，可以静脉注射将脂质体运送到肝脏中实施基因转移，大部分肝脏都可以被转染，还可将脂质体注射皮肤或肌肉。该方法已经成为基因转移治疗研究中的后起之秀。

4. 聚阳离子-DMSO 转染技术

磷酸钙转染技术十分有效而且易于重复，但最适转染条件的范围比较窄，为此又发展出了一种技术条件并不十分苛刻的聚阳离子-DMSO 的转染技术。这种实验程序主要包括两个步骤：先用聚阳离子 Polybrene（Abbott 公司出品的一种聚溴化季铵阳离子的商品名）处理，增加 DNA 对细胞表面的吸附能力；然后再用 25%～30%的 DMSO 短暂处理细胞以增加膜的通透性，提高对 DNA 的捕获数量。按此法测定逆转录病毒 DNA 的转化效率，结果是同 DNA 的加入量成正比，而且不需要加运载 DNA（carrier DNA）就可得到稳定的转化。聚阳离子-DMSO 转染技术的通用性尚未验证，但已知可适用于鸡胚细胞和小鼠成纤维细胞的转化。

三、生理方法

1. 受体介导的基因转移

受体介导的基因转移方法是在质粒 DNA 通过交联剂和某种特异的配体（多肽）之间形成复合体，而这种配体能为细胞表面的受体所识别，之后通过受体介导的细胞内吞途径，将外源基因导入细胞内。

如若将 DNA 在体内运送至肝细胞内，可以选择将 DNA 和能与肝细胞受体特异结合的去唾液酸糖蛋白质偶联，以便通过细胞内吞过程而被摄入，这种 DNA 大部分被肝脏所摄取。应用该方法转移的外源基因在活体内的表达持续时间较短，在评估实际应用前景上还存在一些问题。

2. 胚胎干细胞法

胚胎干细胞是从受精卵开始分裂至 4 个囊胞阶段未分化和具有多种潜能的生殖细胞，能在体外培养，制备转基因动物，以研究基因定向整合或基因剔除及基因功能。

3. 精子载体法

最近发现用精子和 DNA 孵育，可捕获得 DNA。通过受精过程，将外源性基因导入受精卵，大大简化了转基因动物的制备过程。

4. 抗体转染法

利用抗体作为载体介导基因进入表达特定表面抗原细胞的方法为抗体转染（antifection）。

该方法具有简单、安全和适用性广泛等特点。将抗 CD3、CD34 或表面免疫球蛋白的抗体与质粒共价偶联，可自体内或体外转染正常脾 B 细胞或类淋巴细胞。通过抗体介导这一生理过程，基因被传送到特定细胞。该方法可望用于基因治疗研究。

5. 病毒载体介导的基因转移

详见下节。

第三节　病毒载体介导的基因转移

近年来，病毒载体研究已成为基因转移载体研究的主攻方向。病毒作为基因转移的有力工具，主要优点有：①在转化的细胞中传播重组的 DNA 分子，并作为其稳定的遗传成分；②可以将有缺陷的或突变的基因置于病毒调节信号的控制下进行研究；③能将克隆的基因作为病毒微染色体的一部分，并能进行分离；④转移效率较高。

病毒载体主要缺点是对外源基因的最大容纳量只有 2500bp。目前，常用的病毒载体主要有两类，RNA 病毒类和 DNA 病毒类。其中最常用的 RNA 病毒载体是逆转录病毒载体；DNA 病毒载体主要是腺病毒（Ad）载体、腺相关病毒（AAV）载体、单纯疱疹病毒（HSV）载体及多种形式的嵌合病毒载体等。

一、RNA 病毒载体及其介导的基因转移

（一）逆转录病毒

1. 种类

逆转录病毒（retrovirus）是一类含单链 RNA 的动物病毒。它的基因组含有 2 条相同的正链 RNA 分子，包装成二倍体病毒颗粒。除此之外，在其病毒颗粒内部还包含有 tRNA 引物分子、逆转录酶、RNase H 和整合酶等组分。迄今已经在鸟类及哺乳类动物中发现了许多种类的逆转录病毒，例如鸟类的脾坏死病毒（SNV）、鼠类的乳腺肿瘤病毒（MMTV），莫洛尼小鼠白血病病毒（Moloney murine leukemia virus，MoMLV）等，但其中研究得最为详尽的则要数劳斯肉瘤病毒（Rous sarcomavirus，RSV）。RSV 病毒感染了鸡之后，就会诱发产生肿瘤。从感染细胞中分离出来的逆转录病毒的 DNA，是一种有用的动物细胞的基因载体。

逆转录病毒家族可以分成三种不同的类群：它们是泡沫病毒（foamy viruses）类群，例如人泡沫病毒（HFV）；慢病毒（lentivirus）类群，例如人免疫缺损病毒（HIV）Ⅰ型和Ⅱ型，以及绵羊髓鞘脱落病毒（visnavirus）等；致癌病毒（oncovirus）类群。不过需要特别指出的是，并非所有致癌病毒类群的病毒都是致癌的。根据在电子显微镜下观察的病毒颗粒的形态学特征的差异，致癌病毒类群又可区分成 A、B、C、D 4 个类型。

研究表明，逆转录病毒具有许多优点，便于发展作为动物基因克隆载体。第一，就目前所知，在大多数情况下，逆转录病毒的致癌基因（oncogene，onc）都能够在正常的细胞中转录。这种特性说明逆转录病毒有可能是一种天然的转导因子。同时，根据这种特性我们可以在正常细胞中进行操作，将它改建为有用的动物基因的转移载体。第二，逆转录病毒的寄

主范围相当广泛，包括无脊椎动物和脊椎动物，其中有的还能够在人体细胞中生长。当两种不同的逆转录病毒在同一寄主细胞中生长时，就会出现一种 DNA 被另一种外壳蛋白质包装成假型病毒的奇特现象。由于病毒的寄主范围实质上是由其蛋白质外壳决定的，因此使用这种假型病毒便有可能将某些病毒基因组导入正常情况下无法感染的细胞，从而进一步扩大寄主范围。第三，逆转录病毒具有强启动子，克隆此类载体上的外源基因有可能得到有效的表达。有些逆转录病毒，例如鼠乳腺肿瘤病毒，其启动子的活性功能的开启或关闭可以通过实验手段予以控制，因为这种病毒基因的转录是受糖皮质激素诱导的。第四，逆转录病毒不但感染效率高，而且通常还不会招致寄主细胞的死亡，它感染的或转化的动物细胞能够持续许多代，保持正常生长和形成感染性病毒颗粒的能力。因此，有可能利用逆转录病毒作载体，改变动物细胞的基因型，并可遗传到子代细胞。

2. 逆转录病毒生命周期

典型的逆转录病毒的生命周期，是从它的感染颗粒同寄主细胞表面特定的接受器结合，并进入细胞的时候开始的。整个周期可以分为两个不同的时相。在第一时相，首先是在细胞质中的病毒颗粒脱掉蛋白质外壳，释放出的（＋)RNA 在病毒颗粒带进来的逆转录酶的作用下，合成（－)DNA。由此形成的 RNA-DNA 双链分子中的（＋)RNA 链，通过逆转录酶加工作用，被逐渐地降解掉。剩下的（－)DNA 继续指导逆转录酶合成（＋）DNA。最终形成的双链形式的 DNA 分子整合到寄主细胞的染色体基因组上，成为前病毒（provirus)。因此，逆转录酶实际上进行了三种有序的催化反应：RNA 指导的 DNA 合成；RNA-DNA 双链分子中 RNA 链的降解；DNA 指导的 DNA 合成。

在双链形式的 DNA 进入寄主细胞核成为前病毒之后，便开始了逆转录病毒生命周期的第二时相。此时，前病毒 DNA 进行活跃的转录，合成出来的转录本被加工成 mRNA，进而翻译为 2 种成品蛋白质（finished proteins）和 2 种多聚蛋白质（polyproteins)。随后这 2 种多聚蛋白质被切割成 6 种多肽链，其中两条聚集为逆转录酶。在这些蛋白质中，有 4 种形成病毒外壳，另外 2 种成为病毒的被膜（envelope）或叫病毒表面蛋白质。病毒颗粒是在细胞质中组装的，成熟后即通过出芽（budding）的形式从感染细胞中挤压出去，而细胞本身并没有任何损伤，因此可以长时间持续地释放病毒颗粒。

（二）逆转录病毒载体

目前，已经发展出来的以逆转录病毒为基础的动物基因转移载体主要有：需辅助病毒互补的重组的逆转录病毒质粒载体、无需辅助病毒互补的重组的逆转录病毒质粒载体、广寄主范围的逆转录病毒载体，以及逆转录病毒表达载体 4 种类型。下面逐一作简单介绍。

1. 需辅助病毒互补的逆转录病毒质粒载体

需辅助病毒互补的逆转录病毒质粒载体，是 20 世纪 80 年代初发展起来的一种简单的逆转录病毒基因转移系统。其关键的步骤是利用 DNA 体外重组技术，将克隆在大肠杆菌 pBR322 质粒载体上的前病毒 DNA，移去 *gag*、*pol* 和 *env* 等个基因的大部或全部序列，保留下 5′-LTR、3′-LTR 序列以及 PBS-ve、PBS＋ve 和包装位点 psi。然后，将一种作为选择标记的外源基因，例如 *neo* 等插入其中，构成重组的逆转录病毒质粒载体。在选择标记基因插入时，要使它的起始密码子 ATG 恰好是安放在病毒 *gag* 基因的原来位置上。这样，插入的选择标记基因才会在 5′-LTR 指导下转录，并且也才会从正确的 ATG 密码子启动翻译。

经过这样体外操作建成的重组的前病毒 DNA，因为是克隆在 pBR322 质粒载体上，故可以在大肠杆菌细胞中增殖。然后，将纯化出来的 DNA 按标准的转化程序导入适当的受体

细胞，从中筛选出稳定的转化子。这样的转化细胞将能够表达选择标记基因，并合成出重组的逆转录病毒的 RNA 分子。

下一步是要设法将转化细胞产生的重组病毒的 RNA 分子，用一种逆转录病毒的蛋白质外壳包装起来，使之成为具感染性的病毒颗粒。但由于此种带外源选择标记基因的重组病毒基因组是有缺陷的，不能合成自身的蛋白质，故包装所需的全部蛋白质需靠辅助病毒提供。最简单的一种办法是将辅助病毒超感染给转化的细胞。辅助病毒所产生的 Gag 和 Env 蛋白质，能够识别重组病毒 RNA 分子上的包装位点 psi，因此能形成病毒颗粒。而且，这种病毒颗粒也同样含有由辅助病毒合成的逆转录酶，及同 PBS-ve 位点杂交的 tRNA 引物。所以，它能够感染第二种寄主细胞，合成自己的 RNA，最后整合到新寄主细胞的染色体，作为前病毒形式存在下去。当然，如果没有辅助病毒的存在，这种重组前病毒就不可能进一步产生出感染性的病毒颗粒。因此，通常称这种病毒为假病毒，也就是说它具有初始感染的全部必需的蛋白质，但它的基因组却是有缺陷的。此种利用一种病毒的蛋白质包装另一种病毒基因组的方式叫做“假型包装”(pseudotyping)。

2. 无需辅助病毒互补的逆转录病毒质粒载体

应用需辅助病毒互补的逆转录病毒质粒载体转移基因，具有多方面的优点。首先，重组病毒的 RNA 被包装成病毒颗粒之后，就成为一种比质粒载体更加有效的基因转移体系，而且这样的病毒颗粒能够感染具有适当表面接受器的任何细胞，其基因转移成功率几乎可达 100%。其次，这种病毒颗粒在感染了寄主细胞之后，重组 DNA 可以整合到染色体上成为前病毒，而且每个基因组中通常都只有一个拷贝的前病毒，这样就不会出现多重插入的复杂性。

然而从另一方面考虑，这种基因转移系统也存在一些不足之处。由于使用的感染性辅助病毒在感染过程中会同重组病毒竞争细胞表面接受器，从而相对地降低了被重组病毒感染的细胞数量，而且感染性辅助病毒可能是致病的，这就限制了此种体系在整体动物及人体细胞感染方面的使用价值。为了避免这样的麻烦，人们又发展出了不需要辅助病毒互补的重组的逆转录病毒质粒载体系统。这种系统需建立一种特殊的包装细胞株 (packaging cell line)。

包装细胞株的基本特性是：在它的染色体 DNA 的某个位点上整合着一个缺失了 psi 序列的逆转录病毒的前病毒 DNA，即 5′-LTR-*gag-pol-env*-3′-LTR 区段；或是在其染色体 DNA 的两个位点分别整合着缺失了 psi 序列的 5′-LTR-*gag*-3′-LTR 区段和 5′-LTR′-*pol-env*-3′-LTR 区段。因此，在包装细胞中能够组成表达逆转录病毒的全部蛋白质，但由于在它们合成的 mRNA 分子中缺失了 psi 序列，故只能形成中空的病毒外壳。

这里我们用 xF2 细胞为例，说明包装细胞株的构建及重组的逆转录病毒载体 DNA 的包装过程。首先将克隆在大肠杆菌质粒载体 pBR322 上的 MoMLV 原病毒 DNA，缺失掉一个包括 psi 位点在内的长 350bp 的区段。然后，把由此得到的缺陷性的逆转录病毒质粒 pMOV 通过磷酸钙共沉淀法转化给相关的哺乳动物细胞，结果从转化子中筛选到一个其染色体 DNA 上稳定地整合着 pMOV 的包装细胞株 32。

当用缺陷性的重组的逆转录病毒质粒载体，例如 pMLV-gptDNA 转染包装细胞株 32 时，它便会整合到寄主染色体 DNA 上进行转录，合成出具有 psi 序列的逆转录病毒载体 RNA。在包装细胞中，由 pMOV 原病毒 DNA 产生的蛋白质，尽管不能包装自己的病毒 RNA，但对于那些异源病毒 RNA 仍具有包装能力，因此能产生出有感染性的重组的逆转录病毒颗粒。这种病毒可以转化细胞，但不能形成逆转录酶，也不能再形成感染性病毒颗粒，

因此它不会给感染的寄主细胞带来辅助病毒的污染。

3. 广寄主范围的逆转录病毒载体

在 psi2 细胞株中包装形成的假型病毒，同 MoMLV 具有同样的寄主范围。然而令人遗憾的是，MoMLV 的寄主范围却是十分有限的，它只能感染小鼠及其亲缘关系密切的啮齿动物细胞。具有这种类型寄主范围有限的逆转录病毒，属于亲嗜性逆转录病毒（ecotropic retrovirus）。另外一些逆转录病毒能感染与原寄主动物没有亲缘关系的多种动物细胞，它们则属于兼嗜性逆转录病毒（amphotropic retrovirus）。这种广寄主范围的特性使兼嗜性逆转录病毒在发展动物基因转移系统方面具有很大的潜力。但由于鼠类兼嗜性逆转录病毒的分子生物学研究还不够深入，暂时还不能用作基因克隆的载体。因此，根据病毒寄主范围主要取决于病毒表面蛋白质（尤其是 gp70）这一原理，广泛采用兼嗜性逆转录病毒的 *env* 基因取代亲嗜性逆转录病毒 *env* 基因的办法，构建广寄主范围的逆转录病毒载体。其具体操作程序是，先把逆转录病毒的前病毒 DNA 克隆在大肠杆菌质粒 pBR322 上，然后经重组和缺失，形成具有广寄主范围的包装缺陷的病毒质粒。psiAM 和 PAl2 等便是稳定地整合着这种病毒质粒的辅助细胞。实验证明，将带有选择标记和外源基因的重组的逆转录病毒，感染这样的细胞株，便可产生出具有广泛寄主范围（包括啮齿动物、鸟类及人体细胞）的重组的逆转录病毒，而且其产量也很高，每毫升裂解液可达 10^5～10^7 个病毒颗粒。

4. 逆转录病毒表达载体

逆转录病毒的基因组能够承载数个外源基因，最大的插入片段可达 6kb 左右，这样重组的逆转录病毒仍能正常地增殖。而如果在包装之前能通过剪辑作用使基因组长度缩短，那么插入的外源 DNA 则可增加到 20kb。由此可见，逆转录病毒基因组具有相当大的克隆能力。现在，已经发展出许多种逆转录病毒表达载体。例如，有人用 pLPL 载体克隆大鼠生长激素基因，获得了表达。pLPL 载体基因组上带有一个人的 *hprt* 基因作为选择标记，同时只含有一个 *Hpa*I 识别位点。把正常的大鼠生长激素基因的启动区，同一种生长激素 cDNA 融合而成的大鼠生长激素微型基因（mini-gene），插入在 pLPL 载体的 *Hpa*I 位点上，产生出一种逆转录病毒表达载体 pLPGHL。携带有这种表达载体的寄主细胞，能够合成出高效价的感染性的重组体逆转录病毒。再用这种病毒感染大鼠的 $hprt^-$ 细胞，转化的细胞便获得了合成生长激素的功能。而且，应用地塞米松（dexamethasone），可以诱导细胞分泌生长激素。此外，R. D. Cone 等（1987）和 S. Karlsson 等（1987），也用逆转录病毒表达载体实现了人珠蛋白基因的正确表达。

上面所述的例子，插入在逆转录病毒表达载体 pLPGHL 上非选择的外源基因，即大鼠生长激素基因 rGH，是利用自己的启动子 PrGH 进行表达的，因此文献中有时也称这种载体为内部启动子载体（internal promoter vector）。另外一种表达载体，则是利用逆转录病毒 LTR 启动子表达外源基因的，例如 pMSV-hPTH-gpt，它带有人副甲状腺激素基因（hPTH）和 *Eco-gpt* 基因。由于在这种载体中同时插入了两个外源基因，所以有时也称为双基因表达载体（dural expression vector），简称 DE 载体。

有许多事实表明，逆转录病毒表达载体本身的结构因素会影响它在某些细胞中的表达效率。特别是 5′-LTR 序列，在诸如胚胎癌性细胞（embryonal carcinoma cell，EC）等一类未分化的细胞中，对细胞因素的抑制作用十分敏感。还有一些事实表明，同一载体上携带的多重启动子之间会发生竞争作用。例如，有人曾把 HSV-tk 基因及小鼠 β-珠蛋白基因，连同它们自己的启动子一道插入到一种鸟类脾坏死病毒载体上，结果就发生两种启动子竞争表达的

现象。

使用逆转录病毒载体时，还必须考虑到许多方面的其他因素。首先，插入的外源 DNA 不应该有转录终止信号，因为终止信号会阻断形成全长的病毒转录本，而这种转录本则是产生前病毒必不可少的先决条件。其次，如果使用的载体具有 5′剪接位点（splice site），那么十分要紧的是任何插入序列都不应具有也不应该产生一种不恰当的 3′剪接位点，否则将影响插入 DNA 的表达。

综上所述，逆转录病毒载体的特点是：①逆转录病毒感染细胞的效率高，基因转移效率几乎可达到 100%；②病毒可以容纳外源基因 DNA 长度约 7kb；③可将外源基因整合至宿主染色体，外源基因在宿主内稳定表达，并随细胞分裂而传代；④宿主范围广。

该类载体也有其局限性：①逆转录病毒只感染分裂细胞。但人们发现 HIV-1 病毒可以感染 T 淋巴细胞等不分裂细胞，以其改造而来的新型逆转录载体可望能够感染某些非分裂细胞，以扩大其在基因治疗中的应用范围；②病毒滴度低。人们正在研究各种方法来提高病毒的感染效率，但如果使用不同的病毒包装蛋白，如 VSV 的 G 蛋白，包装产生假型病毒，该病毒很容易被浓缩，可提高病毒滴度 1000 倍以上，最高达到 10^9pfu/ml；此外，还可通过改造载体或采用离心方法相对浓缩病毒等以提高病毒的滴度；③要求靶细胞 DNA 合成必须很活跃，因为病毒的感染和插入作用依赖于靶细胞 DNA 的复制；④靶细胞表面需要有特殊的受体；⑤在体内，逆转录病毒易被补体灭活，故目前采用离外（*ex vivo*）途径较多。

由于逆转录病毒作载体的基因导入过程简单、容易操作，是经典的、也是目前常用的基因转移方法，虽然有潜在致癌性，但尚未发现其导致癌症的具体事例。

二、DNA 病毒载体及其介导的基因转移

DNA 病毒载体主要有腺病毒载体、腺病毒相关病毒载体等。对 DNA 病毒载体的研究是当前基因基因治疗研究中的重点领域。

（一）腺病毒载体

腺病毒（AV）是一种线性双链 DNA 病毒，基因组长约 30～50kb。对其结构和功能了解得最清楚的是腺病毒血清型的 2 型和 5 型（Ad2 和 Ad5），这也是作为基因治疗载体中最常见的两个血清型。

腺病毒载体的基因转移原理类似于逆转录病毒（retrovirus，RV）载体，也是通过病毒的反式互补得以实现。经过改造的缺失了 E1 后的病毒无复制能力，只有在能反式提供 E1 基因产物的包装细胞（如 293 细胞）的帮助下才能进行繁殖。293 细胞来源于转染了 Ad5 基因组 DNA 的人胚胎肾细胞，目的基因则可以通过重组插入到 E1 基因的位置，其转录可受 E1 启动子或外源启动子的控制。

AV 是继 RV 之后在基因治疗中广为应用的一种基因转移载体，至目前 AV 已发展到第三代。第一代 AV 缺乏 E1 和 R3，以便插入治疗基因，但该载体的临床试验结果并不理想，主要表现在炎症的产生，以及基因表达的快速衰退，这主要是因为低水平表达的病毒蛋白诱发了宿主的免疫系统反应。于是，第二代 AV 在删除 E1 的基础上，进一步去除了野生型腺病毒的部分结构基因，如 E2、E4 区，并对 293 细胞也做了一些改造，以包装改进后的重组腺病毒。这种载体容量有所增大，同时细胞毒性和免疫原性有所减弱，但仍未能延长目的基因的表达时间，且病毒滴度只有第一代的 1/10～1/100。第三代 AV 有了较大的突破，野生型病毒的结构基因已全部删除，只保留了腺病毒必要的顺式作用元件，即基因组两端的反向

重复序列和包装信号顺序，又称微小腺病毒（mini Ad），被去除的基因功能由辅助病毒通过反式作用来补偿。第三代AV不仅容量大，细胞毒性和免疫原性也大幅减弱，且基因的表达时间也大大延长。

AV之所以成为基因治疗的又一研究热点，主要可归结为以下特点：①腺病毒安全性好，在美国已有多年的临床免疫试验结果证实AV（研究最多的是Ad-5和Ad-2）用于基因治疗是安全的；②与RV相比，AV可携带更长的治疗基因；③AV感染的宿主范围广，可感染分裂和不分裂细胞，拓宽了靶细胞的选择范围，且感染效率最高可达100%；④AV可以在呼吸道及消化道感染，因而可以有多种AV所介导的基因治疗途径；例如静脉注射重组病毒，将重组AV制备成胶囊或药液口服经肠道吸收或喷雾、滴注等，这些方法易于推广，具开发应用价值；⑤AV易于制备、纯化和浓缩，病毒滴度可以高达10^{11}pfu/ml，因而AV介导的基因治疗常采用体内（*in vivo*）方式进行。

AV的主要缺点有：①由于腺病毒可导致人体的免疫反应，外源基因只能作短暂表达，如何避开人体的免疫机制是腺病毒载体设计的重要问题；②腺病毒的广泛感染范围使得局部性应用较为困难，即AV缺乏明确的趋向性。如何修正腺病毒与其受体的作用部分，以提高载体的趋向性，是当今研究的另一个重要方向。

（二）腺病毒相关病毒载体

腺病毒相关病毒（AAV）是目前已知的动物病毒中最简单的一类单链线状DNA病毒，也是迄今为止，所发现的唯一与人类疾病无任何相关性的病毒。目前已发现有5种血清型的AAV，其中对2型（AAV-2）的研究最为深入。AAV-2的基因组由4860bp组成，两端为两个倒转末端重复（inverse terminal repeat，ITR）序列，中间部分主要包括*rep*和*cap*两个基因，分别编码病毒的外壳蛋白和复制蛋白。

AAV载体的构建与逆转录病毒相似，也采用反式互补原理：在AAV基因组中去除*rep*和*cap*基因，插入目的基因及其调控顺序，然后与包装质粒共转染293细胞，包装质粒提供病毒复制、表达和包装所必需的*rep*与*cap*基因，通常在*rep*和*cap*基因顺序的两端加上腺病毒的末端重复顺序，以便于辅助病毒的识别，在同时存在辅助病毒-腺病毒感染时，即可产生带有目的基因的重组腺相关病毒，随后用物理方法如超离心或加热可以将其与腺病毒分离开来。值得注意的一点是为防止野生型AAV的污染，包装质粒和载体质粒应不具同源性。

AAV载体的优点是：①无致病性；②宿主范围广，可以感染分裂和非分裂细胞（如肠、神经细胞等）；③可以整合。AAV中一种B19的病毒能够特异性地整合到宿主第19号染色体上，这样不仅为外源基因的表达提供了一个固定的环境，使目的基因得到长期稳定的表达，也可进一步减少插入突变的可能性。AAV载体的缺点是：①载体容量小，不能超过5kb（其基因组长度）；②包装效率低，一般认为这是由于*rep*和*cap*基因的表达水平较低造成的，而*rep*和*cap*基因的高水平表达有明显的细胞毒效应。为解决这一矛盾，有人采用Tet-on启动子控制*rep*基因的表达，使*rep*在短时间内高效表达，完成包装后即被关闭。

（三）疱疹病毒载体

单纯疱疹病毒（HSV）载体的优点在于：①宿主范围广，尤其对神经细胞具亲和性。除此之外，HSV还能感染肌肉、肿瘤、肝、肺、胰岛细胞等多种细胞，且病毒滴度可达$10^8 \sim 10^9$pfu/ml；②HSV载体可容纳高达50kb的外源DNA，为各种病毒载体之最。

目前主要有两种HSV载体，一种为重组病毒型（recombinant），在HSV非必需区插

入外源基因，而后经同源重组再筛选出重组病毒；另一种为重组质粒型（amplicon），其主要成分为：①在细菌中扩增所必需的复制起始点和抗药性，如 *Amp*；②在真核细胞中扩增必需的复制起始点，如 SV40 的 *ori*；③HSV 病毒的包装信号；④外源基因的表达框。构建的质粒 DNA 可连同 HSV-1 辅助病毒一起导入细胞，质粒 DNA 被串联在一起，包装在病毒颗粒内。这种包装有质粒 DNA 的假型病毒可以感染其他细胞，但其效率较重组病毒低，外源基因的容量也受到限制。

（四）嵌合（杂合）病毒载体

嵌合病毒载体（chimeric viral vector）是近几年刚发展起来的新型病毒载体，其通过分子生物学技术把两种或两种以上的病毒载体结构组合起来，充分发挥各家之长而避其短。该类新型载体的研究正成为基因治疗载体研究的又一重要方面。嵌合载体包括腺病毒/逆转录病毒嵌合载体、腺病毒/腺相关病毒嵌合载体、单纯疱疹病毒/腺相关病毒嵌合载体。

第四节　报告基因的应用

一、报告基因及其应用范围

（一）报告基因

报告基因（reporter gene）是一种编码可被快速地测定、检测的蛋白质或酶的基因。报告基因常被用来判断外源基因是否已经成功地导入宿主细胞，以及检测基因的表达活性。在报告基因的基础上发展起来的基因报告体系（genetic reporter system），如把它的编码序列和基因表达调节序列相融合形成嵌合基因，或与其他目的基因相融合，在调控序列控制下进行表达，从而利用它的表达产物来标定目的基因的表达调控，已被广泛地应用于基因工程特别是真核基因表达与调节方面的研究。

在基因工程的研究中，当人们分离到了目的基因之后，一个重要的工作便是对其功能进行分析与鉴定。因此，必须把外源目的基因有效地转化到宿主的细胞或组织中去。鉴于在转化的群体中，只有极少数的细胞被转化，故为了有效地筛选已经获得了外源基因的转化子细胞，并鉴定外源目的基因在宿主中的功能作用，就必须使用适当的报告基因。只用通过分析报告基因的表达产物 mRNA 和蛋白质，才可检测出在转基因细胞中报告基因的活性。

作为一种理想的报告基因，在遗传选择和筛选检测方面必须具备以下几个条件：

① 已被克隆和全序列已测定；

② 在被转染的细胞中不存在该表达产物，即无背景；无相似的内源性表达产物或易于和内源性基因区别；对转化的细胞无毒害，而且该基因的表达必须不改变受体细胞或生物的生理活动；

③ 其表达产物能进行定量测定，同时具有灵敏度高、重复性好、操作快速、简便、经济等特点；

④ 报告分子的分析结果应具有很宽的线性范围，以便于分析启动子活性的变化。

一般在实验设计中，是将报告基因的编码区与位于其上游的目的基因融合，并置于目的

基因启动子的控制之下。可见，报告基因实质上起到了判断目的基因是否表达的标记基因的作用，故报告基因有时也可叫做选择标记基因（selectable marker gene）。

各类报告系统在不同程度上符合上述标准。每一报告基因及其相应的分析系统都各具特色及局限性。人们通常是根据特殊的实验要求，决定选用何种报告基因。例如，报告分子可能在细胞内表达，或从细胞中分泌出来在培养液中检测到，或因有内源性活性而导致高背景而受到限制。若报告基因表达的产物会干扰细胞的正常功能，这样的报告基因就不适用于转基因动物的培育。此外，也有一些学者认为，报告基因及其产物的存在，有可能使转基因动物受到“污染”。

（二）报告基因的应用

在分子生物学及基因工程的研究中，报告基因除了用于检测外源基因的转化效果和筛选转基因细胞之外，对于真核基因表达调节机理的分析，亦是相当有用的。在诸如重组事件、基因导向（gene targeting）、RNA 加工以及信号传导途径等研究工作中，报告基因都有着广泛的用途。

1. 启动子的检测

报告基因可用来分析控制基因表达的顺式作用元件。通过构建具有特定报告基因而没有启动子的报告载体（reporter vector），可方便地对有关的启动子及增强子元件的功能和特性，做出准确有效的鉴定。

在这种类型的鉴定工作中，人们是将待检测的启动子克隆在报告基因的上游，而增强子由于其功能作用与位置无关，故既可克隆在报告基因的上游，亦可克隆在报告基因的下游。于是，在无启动子的载体中报告基因的表达活性，事实上是由待测启动子及增强子决定的。所以，根据对细胞中报告基因产物的分析，就可确定待测启动子或增强子的功能与特性。应用这种报告基因技术，已经成功地鉴定出了一批在特定类型的细胞、组织或特定发育阶段表达的启动子和增强子。

2. 反式作用因子的检测

在基因表达调控的研究中，经常要涉及到顺式作用元件（cis-acting element）和反式作用因子（trans-acting factor）的概念。所谓顺式作用元件，指的是能够影响同一条 DNA 序列或相邻 DNA 序列活性的特定的 DNA 区段。例如，启动子中与转录因子结合的 DNA 区段，便属于顺式作用元件，它可影响该基因或同一染色体上另一个基因的表达活性。而反式作用因子则是指能够影响基因组中同一条染色体上相同基因或另一条染色体上另一个基因表达活性的基因编码产物（通常是蛋白质），它通过与顺式作用元件之间的相互作用而发挥对基因表达活性的调节作用。

为了检测反式作用因子，可将含特定启动子和报告基因的表达载体，同另一个表达反式作用蛋白质的克隆载体，共转化到同一寄主细胞。如果所表达的反式作用蛋白质能够同表达载体中的启动子结合，就会促使报告基因的表达活性上升。据此，便可对反式作用因子进行检测，进而在此基础上，运用转录因子分离纯化技术，分离并克隆反式作用因子。

3. 相关蛋白质的分离

S. Fields 等发展的酵母双杂交技术，根据报告基因的表达活性，能够在活体内鉴定一对相互作用的蛋白质，并用来分离与已知的蛋白质相互作用的另一种待分离的未知蛋白质。在酵母双杂交系统中，这两种相互作用蛋白质，一种是已知的叫做诱饵蛋白质（X），它与 DNA 结合域 BD 融合，构成 BD-X 质粒载体；而另一种待分离的未知蛋白质（Y）则是同转

录激活区 AD 融合，构成 AD-Y 质粒载体。将这两种质粒载体共转化酵母细胞，所表达的这两种融合蛋白质，通过其 X 蛋白和 Y 蛋白之间的相互作用，致使 BD 和 AD 在空间上接近，从而形成有功能的转录因子，激活下游的报告基因进行表达。之后，检测报告基因的表达产物及其活性，筛选转化体。

由此可见，报告基因的活性为筛选阳性菌落提供了方便。得到阳性菌落之后，则可进一步检测已知蛋白质间的相互作用、研究蛋白质相互作用的结构域或重要的活性位点，以及从 cDNA 文库中筛选与已知蛋白质 X 相互作用的未知蛋白质 Y 的基因编码序列等。

4. 基因工程细胞系的构建

报告基因的另一个重要用途是构建基因工程细胞系（engineered cell lines）。我们知道，如果在转化的载体上除了通常具有的抗生素抗性基因和报告基因之外，再加上一个选择标记基因，那么在转化到寄主细胞之后，就可方便地挑选出在其染色体基因组中已经整合了报告基因的稳定的细胞系，这就是所谓的基因工程细胞系。它可以用来进行药物筛选（drug screening）、外源药物的检测以及刺激基因表达等多种用途。

二、常见的几种报告基因及其应用

目前研究工作者已经从大量的基因中挑选出了若干种有用的报告基因，其中效果最好的有荧光素酶（luciferase，LUC）基因、β-葡萄糖醛酸糖苷酶（β-glucuronidase，GUS）基因和氯霉素乙酰转移酶（chloramphenicol acetyltransferase，CAT）基因等。最近，从多管水母（*Aequorea victoria*）中分离出来的绿色荧光蛋白质（green fluorescent protein，GFP）的编码基因也是一种在基因工程中有用的报告基因。

根据报告基因系统表达产物的特点，可分为体外报告分子的分析和体内报告分子分析两种。

体外报告分子的分析是指用含有报告分子的细胞或组织裂解液，或（对于分泌出来的报告分子）被转染的细胞的培养液来进行报告分子定量的方法。这些分析方法是通过酶学或免疫学方法直接定量报告分子，从而间接估计报告载体编码的报告分子的转录活性。这样的分析方法与下述的体内报告系统不同，能获得适合于对比研究不同的启动子和增强子的定量数据，并且，对细胞类型的要求也更容易达到。采用体外报告分子分析的报告基因有编码 GUS、人生长激素（hGH）、分泌型的碱性磷酸酶（SEAP）、β-半乳糖苷酶、萤火虫荧光素酶、氯霉素乙酰转移酶（CAT）等的基因。

体内报告分子分析可定义为在活细胞或组织（如转基因动物）中，或在已经过组织化学固定的细胞或组织中所进行的分析。本方法提供的数据比体外报告系统的定量性较差，但可为启动子/增强子的细胞特异性及特异转录因子的组织分布提供重要信息。由于该方法的关键在于将报告分子固定在其产生的部位，因此，用可析出沉淀的底物或能发光或产生荧光的底物来进行报告分子的分析是十分合适的。相反，分泌型蛋白如 SEAP 及 hGH 的编码基因则较少用作体内报告基因。采用体外报告分子分析的报告基因有 β-半乳糖苷酶、萤火虫荧光素酶、绿色荧光蛋白（GFP）等的编码基因。

（一）荧光素酶

自然界中广泛分布着生物发光有机体，其中包括细菌、真菌、鱼、昆虫等，在这些生物发光有机体中催化生物发光反应的各种酶都称为荧光素酶（luciferases，LUC），底物则命名为荧光素（luciferin）。虽然来自不同物种的荧光素酶及底物存在有很大的差异，但有一

个共同点，即在生物发光反应体系中均需发生氧化反应，该反应放出黄绿光（发射峰为560nm），可为 Luminometer 或标准的液闪仪所检测。

1. 应用

自 1986～1987 年首次被当作报告基因使用以来，荧光素酶已是目前运用最广泛的报告基因之一。迄今为止，被深入研究的昆虫类荧光素酶分离自北美洲萤火虫（*Photinus pyralis*）。与其他不同来源的荧光素酶序列相比较，它们只有 25%的序列在进化中高度保守；因此从进化的角度来看，此类荧光素酶无相关性。1987 年，Jeffrey 等构建了全长的、无内含子的、能插入哺乳动物细胞中瞬时及稳定表达的荧光素酶 cDNA 序列，并用来作为报告基因，运用化学发光等方法显示其活性，从而间接反映待测 DNA 序列的活性强度，如监测功能性启动子或增强子等调控元件的活性、插入及定位等，也可用来监测转染效率，同时还可作为监测重组体启动子和增强子活性的内对照；此外，1996 年新发展了一种双报告（dual receptor）技术，基于萤火虫（firefly）和水母（renilla）的荧光素酶有不同的进化起源，酶结构不同，对底物有不同的要求，可以分别催化甲虫萤光素和腔肠动物素（coelenterazine），针对这一特性，在一个试管中将萤火虫与水母起源的荧光素酶测方法结合起来。其中水母来源的荧光素酶（renilla Luc）作为对照报告基因；可在一个样品中连续检测两种有相似检测速率、灵敏度、线性范围的报告基因。因此，样品只需裂解后处理一次，不需分成几份。

2. 优缺点

在实践中，该报告基因系统具有省时、便宜、灵敏度高（0.02pg，约 250000 个酶分子即可被检测到），较 CAT 检测法灵敏度高 30～1000 倍、无需使用放射性同位素、无需特殊设备的优点。并且，由于荧光素酶能够扩散通过哺乳动物的细胞膜，因此可以不损害细胞，直接测量细胞内的荧光素酶的活性。

人们也广泛研究了实时荧光素酶报告基因反应的动力学，发现反应在 0.3s 内达到峰值，其后有一持续时间较长的、强度较低的时期。在适当的条件下，发光峰的高度与反应液中的酶浓度成一定的比例。理论上，由于哺乳动物细胞中无内源性的荧光素酶酶活性，因此该反应检测下限取决于仪器的信噪比，上限取决于电极处理系统，通常的线性范围为 0.02～10pg 荧光素酶。

3. 体内分析的相关问题

近来的研究显示，通过应用可穿过质膜的可溶性荧光素酶底物，可在活细胞内检测到萤火虫荧光素酶的活性。这些化合物一般是不带电荷的荧光素酯类，很容易穿过脂质双层膜。一旦进入细胞中，这些底物即被内源性的酯酶转变成萤火虫荧光素，后者即成为报告基因所编码的荧光素酶的底物。有一种可光解的“笼型”荧光素化合物，据认为更容易进入细胞。这种化合物能被内源性的酯酶或可见光水解而产生游离的荧光素。以其现有的形式，体内的荧光素酶分析方法还欠灵敏，并且缺乏 GFP 所具有的进行实时分析的能力。通过开发其他有更高的量子产率的荧光素酶底物，以及更复杂的检测仪器，荧光素酶有可能成为一种体内分析用的报告系统。

（二）绿色荧光蛋白

绿色荧光蛋白（green fluorescent protein。GFP）是一类存在于水母、水螅和珊瑚等腔肠动物体内的生物发光蛋白。当受到紫外或蓝光激发时，绿色荧光蛋白发射绿色荧光。其独特之处在于：它产生的荧光持续时间长，无需固定底物或辅因子，发色团是其蛋白质一级序列固有的。来源于维多利亚水母（Aequorea victoria）的绿色荧光蛋白 cDNA 开放阅读框架

长度约 740bp，由 238 个氨基酸残基组成，分子质量为 27kD 的单体蛋白质。水母的 GFP 基因是目前唯一已被克隆的 GFP 基因。

1. 原理

绿色荧光蛋白不仅存在于水母中，在其他一些腔肠动物中，如薮枝（Obelia）珊瑚虫也有发现。它们的发光过程分为两类，分别以水母（Aequorea）和珊瑚虫（Renilla）的发光途径为代表。两者的发光过程都需要 Ca^{2+} 结合到水母发光蛋白上，使荧光素被氧化而释放出蓝光，继而进一步激发绿色荧光蛋白发出绿色荧光；Aequorea 在接受了 Ca^{2+} 激活的发光蛋白水母素的能量后可在体内产生荧光。这个过程的发生不需要底物或辅助因子，而是通过两个蛋白之间的能量直接转移进行的。而 Renilla 发光是由荧光素酶氧化 LBP（荧光素结合蛋白）结合的荧光素产生蓝光并激发绿色荧光蛋白产生的。此外，还有一些发光生物，如水母 *Beroe*、*Mnemiopsis* 中也有 Ca^{2+} 激活的发光蛋白，但它们没有绿色荧光蛋白，因而只能发出蓝光。

全长的 GFP 对于发射荧光似乎是必要的；绿色荧光蛋白的发光基团是由 Ser65、Tyr66 和 Gly67 三个氨基酸经过环化和 Tyr66 氧化（脱氢）后形成的一个咪唑环。其荧光发生与发光基团的周围环境有关，真正发光的是咪唑环上的酚阴离子。

GFP 有两个吸收峰，主峰在 395nm，另有一小峰在 470nm，这两处波长的任何一种光激发都能使绿色荧光蛋白产生 508nm 的绿色荧光。转染后的细胞可在荧光显微镜或流式细胞仪（FACS）中直接观察基因的表达。GFP 的吸收及发射波长与荧光素相似，因此，用来检测荧光素的各种条件也适合于 GFP。

目前 CLONTECH 公司可提供的绿色荧光蛋白载体种类，如提供绿色荧光蛋白或绿色荧光蛋白变化基团的载体、用于观察启动子或增强子活性的转录报告载体、用于异源蛋白融合的 N 端绿色荧光蛋白载体、用于异源蛋白融合的 C 端绿色荧光蛋白载体、双蛋白表达载体、亚细胞定位和共转染标记载体等。

2. 应用

（1）基因表达的检测　主要包括两个方面：①测定质粒的转染效率和整合效率。将绿色荧光蛋白表达质粒转入不同类型的细胞瞬时表达，可以比较不同细胞间转染效率的差异；为了测定 DNA 的整合效率，也可计数稳定表达绿色荧光蛋白的细胞。但在这方面的应用中，绿色荧光蛋白最大的缺点就是没有放大作用，它不能像酶一样能通过催化无数的底物分子而将信号放大。所以，一般需强启动子以驱动绿色荧光蛋白基因在细胞内足量的表达。也可用亚细胞分辨率的显微成像系统检测基因产物，靶入的基因被限制于一个细胞器内，绿色荧光蛋白的浓度则相对提高了许多倍。②在基因工程中作为筛选细胞的标记。将 GFP 与目的基因融合后转入细胞中，可通过荧光显微镜或流式细胞仪筛选产生荧光的细胞，也就找到了表达目的基因的细胞，用于进一步的研究。

（2）转基因动物研究中的报告分子　目前将绿色荧光蛋白应用于动物成体中的研究还刚刚起步；主要是作为转基因动物研究中的报告分子，用来标记外源基因在导入动物胚胎或卵后在成体中的表达。在转基因鼠的研究中由于传统的 PCR 和杂交技术用于检测基因的整合时较为繁琐，使用绿色荧光蛋白可以简便有效确定转基因鼠，以便更快地获得需要的品系。

（3）在细胞生物学中作为荧光标记用于细胞动态的研究　由于绿色荧光蛋白分子较小，能穿过细胞膜与核膜，因而可用作蛋白 C 端或 N 端的荧光标记，用于细胞动态的研究。但其两端只能去除几个氨基酸才不致影响荧光发生。将绿色荧光蛋白与各种细胞器内的蛋白融

合，就能将该细胞器标记为绿色，便于在显微镜下观察其运动过程，如细胞骨架、细胞分泌途径、原生质膜、线粒体和过氧化物酶体的研究。

（4）活细胞内蛋白的定位研究　单一的绿色荧光蛋白遍布于胞核和细胞质内，且不影响其他蛋白质在细胞内的定位，因此可用于示踪蛋白质在活细胞内的位置，提供了一个观察活细胞内蛋白质定位的崭新途径。这对于胞周期乃至信号传导的研究无疑是一个巨大的帮助。

（5）GFP 在 Western blotting 中的应用　将蛋白 A(PA) 用绿色荧光蛋白标记后再用于 Western blotting，这样在 395nm 紫外光照射下即可检测出抗体-PA-GFP 复合物。它具有以下优点：只需紫外光激发；操作简单快速；产生的荧光稳定，至少可持续 30min；表达 PA-GFP 的大肠杆菌的细胞提取物可直接用于 Western blotting；勿需纯化。

（6）绿色荧光蛋白在肿瘤研究中的应用　绿色荧光蛋白还广泛用于肿瘤的研究，如肿瘤细胞的转移示踪、癌基因与抑癌基因的间接机制研究。S. Kagawa 等为了研究 TRAIL 基因的旁路效应，将 TRAIL 基因与绿色荧光蛋白相连后转染肿瘤细胞和正常细胞，检测 TRAIL 基因对肿瘤细胞和正常细胞存活影响，从而探讨该基因在肿瘤细胞中的作用机理。

3. 优缺点

① 水母绿色荧光蛋白的荧光很稳定，无种属限制。已在网柄菌、大肠杆菌、酵母及多种动植物细胞中表达成功并产生荧光。

② 绿色荧光蛋白不需要反应底物与辅助因子。不同于以往常用的 *lacZ*、*LUC*、*CAT*、*GUS* 等报告基因，尤其适用于体内的即时（real time）检测。

③ 研究已证明与绿色荧光蛋白的 N 端和 C 端融合的蛋白质具有绿色荧光蛋白的荧光性质和配体蛋白质的生理功能。

④ 即使在用甲醛或戊二醛固定的细胞中，绿色荧光蛋白的荧光也不受影响，适用于与其他荧光试剂同时进行双标实验。

另外，绿色荧光蛋白荧光强度高、检测方便，可用荧光显微镜、流式细胞计或显微成像技术方便地在活细胞中检测；而对活细胞无伤害。绿色荧光蛋白肽链中一些特定氨基酸的替代可使之产生不同颜色的光；例如 S65T-GFP、GFPh、EGFP、BFP、RSG-FP4、S147P、S65C，这些突变体可适应于不同研究的需要 。

由于这些优点；绿色荧光蛋白已成为监测体内基因表达及细胞内蛋白质原位定位的十分重要的报告分子。

但野生型绿色荧光蛋白也有一些缺点，如用蓝光激发时在 470nm 处的荧光强度较低、通过蛋白质折叠产生荧光的过程明显滞后、复杂的光异构化、蛋白质折叠极易受温度影响以及其基因在某些种属的哺乳动物和植物细胞中的表达较低等，对绿色荧光蛋白作为报告分子极为不利。为改善其性能，已成功地筛选出一系列更适于作为报告分子的绿色荧光蛋白变种，为它的开发及应用带来更加广阔的前景。但是绿色荧光蛋白更精细的结构、发光机制，以及它的合成、加工过程尚需更进一步的研究。另外，作为一种检测手段，绿色荧光蛋白应用于定量研究还刚起步。

将来，绿色荧光蛋白的应用可望在三个方面有新的突破：一是继续寻找荧光强度更大、适应范围更广的突变体，尤其是可以发出红色荧光的突变体；二是荧光中介的能量转化现象的发现和证实为研究细胞内蛋白与蛋白的相互作用提供了新的手段；三是绿色荧光蛋白在体外的应用，利用它的荧光特性与其他一些手段相结合，必然可以产生许多快速简便且灵敏度高的检测手段。

（三）氯霉素乙酰转移酶（CAT）

氯霉素乙酰转移酶符合上述许多作为遗传报告分子的标准。该酶催化乙酰辅酶 A 的乙酰基转移至氯霉素。由于 CAT 是一种原核生物来源的酶，因而在许多真核生物细胞中都很少有竞争活性，使对报告分子的分析可达到很高的信噪比。然而，许多 CAT 分析方法都需要价钱相对昂贵的放射性底物，并且分析操作很费时，其灵敏性比新近发展起来的非同位素报告系统差。CAT 相当稳定（在哺乳动物细胞中的半寿期为 50h），使之十分合适作为短暂转染实验中的报告分子，但对于稳定表达分析则不太理想。而且，CAT 所表现的稳定性在可诱导系统中（如可诱导的增强子）也是不利的，报告分子的累积可能掩盖由于加入效应物而引起的诱导作用。对于这种情况，具有较短半寿期的报告分子如萤火虫荧光素酶则更理想。

最常见的 CAT 分析方法基于 ^{14}C-标记的氯霉素与从被转染细胞制备的细胞裂解物一起温育时发生的反应，在硅化的玻璃板上进行薄层色谱分离乙酰化和非乙酰化氯霉素，将玻璃板对 X 光片进行曝光，从而定性估计 CAT 的活性。为了获得更量化的数据，可从玻璃板上刮下相应于单乙酰化及二乙酰化形式的氯霉素位置的斑点，置于液闪计数仪中进行计数。这种方法在有多个样品时显得特别烦琐，而且由于出现两种产物而难以解释结果。现在发展的差异提取技术可以定量分离氯霉素与其乙酰化形式的衍生物，从而不需要进行费时的色谱分离。在这种方法中，乙酰辅酶 A 被放射标记，^{3}H 可代替 ^{14}C，使分析在某种程度上危害更小，价钱更低廉。选用另一种溶剂系统可改进信噪比而提高灵敏度。

有两种定量分析 CAT 的非同位素检测方法。一种是在 96 孔微量滴定板上进行的 CAT ELISA 分析，可直接用抗 CAT 的抗体（Boehringer Mannheim 公司）对 CAT 进行定量。这种方法的长处在于直接测定 CAT 蛋白水平而不是 CAT 的活性。一种非同位素的 CAT 活性测定方法也发展起来，该方法采用了一个经单乙酰化反应而生成的氯霉素衍生物（来源于 Molecular Probes 公司）。由于用该底物只能产生单一的产物，因而较用氯霉素作底物定量分析 CAT 更可靠。不仅如此，这种方法比用同位素底物的方法更灵敏。并且具有更宽的线性范围。

（四）β-半乳糖苷酶

来源个大肠杆菌的 *lacZ* 基因编码 β-半乳糖苷酶，是用途最广泛的遗传报告分子之一，有各种不同的底物可供进行体外及体内分析。该酶催化各种 β-半乳糖苷类物质的水解，包括几种专为不同分析方法而特制的底物。除了用作未知特性的顺式调控序列的报告分子外，在确定的启动子控制下的 β-半乳糖苷酶的表达也常用作校正其他报告分子分析结果的内对照。β-半乳糖苷酶对校正 CAT 及萤火虫荧光素酶尤其有用，因为这些报告分子分析方法中所制备的细胞裂解物同样适于用来检测 β-半乳糖苷酶的活性。用邻硝基苯-β-D-半乳呋喃糖苷（ONPG）进行比色分析及 4-甲基荧光素-β-D-半乳糖苷（MUG）进行荧光分析的方法都已发展起来，但这些方法的应用很有限，主要是由于其灵敏度低。化学发光的 1,2-二氧杂环丁烷（1,2-dioxetane）底物用了 β-半乳糖苷酶的检测，增加分析的灵敏度及扩大检测的线性范围，大大促进了 *lacZ* 作为转录报告基因的应用。当用发光计来检测化学发光信号时，该分析办法较比色法要灵敏 50000 倍以上。采用减小真核细胞的 β-半乳糖苷酶引起的内源性酶活性的分析条件也可以提高灵敏度。

在原核及真核细胞中、在组织切片及完整的胚胎中体内 β-半乳糖苷酶报告分子的水平用可沉淀性底物 X-gal 可进行检测。与 X-gal 的反应产生一种深蓝色，许多情况下相对于背

景很容易检测到。用 X-gal 进行组织化学染色需要在酶学分析之前对细胞或组织进行固定。检测全胚胎中 *lacZ* 的表达，对于分析发育早期的组织特异性基因表达特别有用。在活体培养的细胞中检测 β-半乳糖苷酶的活性，可用底物荧光素-β-D-半乳吡喃糖苷（FDG）进行。FDG 在低渗条件下加入而被送入活细胞内，被 β-半乳糖苷酶切割后因反应产物的疏水结构而被“困”在细胞内。表达该种报告分子的细胞即通过底物代谢物的荧光成分发出的荧光而被检测。

另一种体内表达 β-半乳糖苷酶的重要应用是所谓的“增强子捕捉”研究，其中报告基因整合到转基因生物的染色体内，其位置效应被用来定位增强子序列。有一种方法是将一个直接由 *lacZ* 基因上游克隆得到的弱基础启动子导入小鼠卵细胞，并整合进小鼠基因组中。从产生的成年动物的子代中可获得稳定的转基因动物。应选择足够弱的启动子，使只有当启动子与整合位点附近的某一增强子临近时才能表达 β-半乳糖苷酶基因。转基因动物品系中 β-半乳糖苷酶的表达形式可间接地反映附近增强子的组织特异性及强度。这种类型的实验利用了报告基因的表达来识别新基因及其旁侧的调控序列，它们在早期发育过程中的表达谱式可能会有应用价值。

（五）分泌型的碱性磷酸酶（SEAP）

分泌型的碱性磷酸酶（SEAP）与上述的报告分子有所不同，SEAP 可从被转染的细胞中分泌出来，因而可用少量培养液进行分析。SEAP 编码人胎盘碱性磷酸酶（PLAP）的一个截短的形式。它缺乏一个重要的膜锚定区域，因此能让蛋白有效地从被转染的细胞中分泌出来。培养液中检测到的 SEAP 活性水平与细胞内 SEAP mRNA 及其蛋白质的浓度变化直接成正比。

SEAP 的不寻常的特性是具有极大的热稳定性及对磷酸酶抑制剂 L-高精氨酸有抗性。因此，可用 65℃预处理样品及加抑制剂共温育而除去内源性碱性磷酸酶活性。SEAP 的这种分泌特性对其作为遗传报告基因具有以下几个优点：①不需要制备细胞裂解物；②通过反复收集同一培养细胞的培养液很容易进行基因表达的动力学研究；③被转染的细胞在检测 SEAP 时不受影响，并保持完好以备进一步研究用；④培养液中几乎没有由内源性碱性磷酸酶活性引起的背景；⑤样品的收集和分析能在 96 孔微量滴定板上自动进行。

最早的 SEAP 分析方法是用碱性磷酸酶的底物对硝基苯磷酸进行的一种比色分析法。这种方法快速、操作简单，且十分便宜，但是其灵敏度差、动力学线性范围窄。以 D-荧光素-*O*-磷酸盐的水解为基础的 SEAP 两步生物发光测定法能提高灵敏度。SEAP 催化的去磷酸化反应产生游离的荧光素，后者作为萤火虫荧光素酶的底物。这种方法的灵敏度约等于分析荧光素酶的传统生物发光分析法。最灵敏的 SEAP 分析方法是用化学发光的碱性磷酸酶底物如 1,2-二氧杂环丁烷（CSPD）。CSPD 的去磷酸化产生一种持续的“增长”型荧光，可持续 60min 之久，易于用发光计或闪烁计数仪检测。SEAP 的化学发光分析不仅提高了灵敏度，而且还大大增加了检测的动力学线性范围。

（六）人生长激素（hGH）

正常情况下，人生长激素仅由垂体前叶的促生长细胞分泌。hGH 表达范围的局限性使其成为适于大多数哺乳动物细胞的颇具吸引力的报告基因。hGH 报告分子有许多与 SEAP 相同的优点，但由于该方法的灵敏度相对较低，并需要用有害的放射免疫分析方法来定量 hGH 而不够理想。hGH 的测定被用来作为内部对照，以校正转染效率；同样，hGH 用来校正实验性 SEAP 报告基因构建体的表达也非常适宜。

（七）β-葡糖醛酸酶（GUS）

微生物的β-葡糖醛酸酶（GUS）基因是用于研究植物基因表达的主要报告基因。GUS在植物及哺乳动物细胞中都可以用作报告基因，但在高等植物中特别有用，因为在大多数种类的植物中都不含内源性的GUS活性。用GUS转化的高等植物能健康生长，正常发育。同β-半乳糖苷酶一样，GUS作为报告基因的主要优点之一是有多种分析该酶的方法。用各种β-葡糖醛酸类物质作底物发展了几种不同的比色分析方法。其中，最普遍的是用X-Gluc，这种底物也能用来对表达GUS活性的组织和细胞进行组织化学染色。一种更为灵敏的GUS荧光分析方法用4-MUG作为底物。一种GUS化学发光分析方法也已发展起来，它很类似于用来定量β-半乳糖苷酶的方法，其底物是金刚烷基1,2-二氧杂环丁烷芳基葡糖醛酸。这种分析方法比用4-MUG为底物的荧光分析方法的灵敏度大约高100倍，并且有很宽的动力学线性范围。GUS除了用作报告基因外，GUS融合蛋白还被广泛用于研究植物及动物蛋白定位及转运。

（屈　伸　王宇哲）

第七章　蛋白质表达技术

蛋白质表达技术主要用于生产体内有重要价值的蛋白质产品，它是通过将外源基因与表达载体构成的DNA重组体导入受体细胞中进行高效表达来实现的，包括外源基因的克隆、转录、翻译、加工、分离纯化等过程。蛋白质表达系统可以分为原核表达系统和真核表达系统两大类。

原核表达系统有大肠杆菌表达系统、芽孢杆菌表达系统、链霉菌表达系统。真核表达系统有酵母表达系统、昆虫表达系统、哺乳动物细胞表达系统和植物细胞表达系统。根据目的蛋白质的情况选择合适的表达系统。一般情况下，常用的表达系统主要是大肠杆菌表达系统和哺乳动物细胞表达系统。

蛋白质表达的目的之一是获得重组蛋白质，因此外源基因表达产物在分离纯化前必须确定宿主细胞是否表达目的蛋白质，以及表达量的高低。这需要对重组蛋白质进行定量分析和纯度分析，然后确定分离纯化的过程。

第一节　外源基因在原核细胞中的表达

一、大肠杆菌表达系统

大肠杆菌表达系统的发展历史可追溯到二十多年前。1976年Struhl证明了外源基因在大肠杆菌中可以实现有功能的活性表达，1980年Guarante以质粒、乳糖操纵子为基础构建了第一个大肠杆菌表达系统。

与其他表达系统相比，大肠杆菌表达系统具有遗传背景清楚、目标基因表达水平高、培养周期短、抗污染能力强等特点，在基因表达技术中占有重要的地位，是分子生物学研究和生物技术产业化发展进程中的重要工具。

（一）大肠杆菌表达载体的表达元件

表达载体（expressing vector）是用来在受体细胞中表达（转录和翻译）外源基因的载体，是外源基因在大肠杆菌中表达所不可缺少的重要工具。表达载体除具有克隆载体所具有的性质以外，还带有表达元件——转录和翻译所必需的DNA序列。

大肠杆菌表达载体要求：①有一个强的启动子及其两侧的调控序列；②应有SD序列，且SD序列与起始密码子ATG之间要有合适的距离；③在克隆基因与启动子之间有正确的阅读框架等；④外源基因下游应加入不依赖ρ因子的转录终止区。

1. 启动子

启动子是DNA链上一段能与RNA聚合酶结合并能起始转录的序列，其大小在20～300bp范围内，具有转录目标基因mRNA的功能。启动子是基因表达不可缺少的重要元件。

原核生物的启动子通常具有一些高度保守的核苷酸序列，如 Pribnow box、−35 区等(图 7-1)。

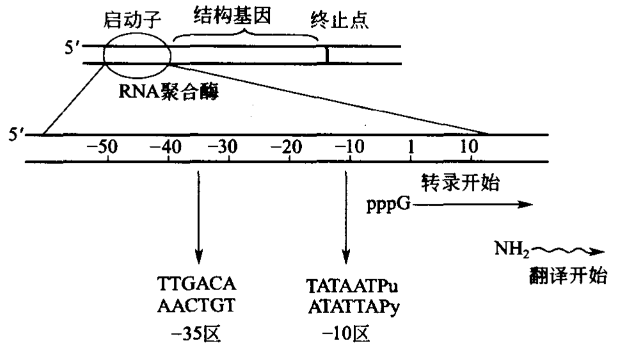


图 7-1 大肠杆菌启动子示意图

Pribnow box 位于转录起始点上游 10bp 处，一般由 6～8bp 组成，富含 A/T，又称为−10 区；−35 区位于转录起始位点上游 35bp 处，一般由 10bp 组成。

一般认为，大肠杆菌 RNA 聚合酶识别并结合启动子，−35 区和 RNA 聚合酶 σ 亚基结合，−10 区和 RNA 聚合酶的核心酶结合。在转录起始位点附近，DNA 被解旋形成单链，RNA 聚合酶使 2 个游离的核苷三磷酸反应形成第一个磷酸二酯键，以后 RNA 聚合酶向前推进，使新生的 RNA 链不断延长。

在不同的细菌中，这些保守序列略有不同，但功能相近，均具有结合 RNA 聚合酶并启动 RNA 转录的能力（图 7-2）。可是，保守序列中的这种碱基的变化会影响 RNA 聚合酶识别和与其结合的能力，从而影响转录水平的高低，同时−35 区与−10 区的距离以及−10 区与转录起始核苷酸的距离都会影响转录水平的高低，所以启动子又分为强启动子和弱启动子。由于 RNA 聚合酶和启动子结合的过程还涉及其他的蛋白因子的作用，所以启动子又可分为组成型和调控型。在大肠杆菌表达系统中，大部分表达载体采用调控型启动子控制目的基因的表达。这对于建立一个高效、实用的表达系统是必不可少的。

	−35 区		−10 区		+1	
trp	T T G A C A	$\cdots N_{17}\cdots$	T T A A C T	$\cdot N_7 \cdot$	A	···
$tRNA^{trp}$	T T T A C A	$\cdots N_{16}\cdots$	T A T G A T	$\cdot N_7 \cdot$	A	···
lac	T T T A C A	$\cdots N_{17}\cdots$	T A T G T T	$\cdot N_6 \cdot$	A	···
recA	T T G A T A	$\cdots N_{16}\cdots$	T A T A A T	$\cdot N_7 \cdot$	A	···
ara	C T G A C G	$\cdots N_{18}\cdots$	T A C T G T	$\cdot N_6 \cdot$	A	···
最大一致性	T T G A C A		T A T A A T			
$x/45$	38 37 36 37 29 28		40 41 25 29 30 44			

图 7-2 几种原核生物基因启动子的保守性序列

由于细菌 RNA 聚合酶不能识别真核基因的启动子，因此大肠杆菌表达载体所应用的启动子必须是原核启动子。将外源基因克隆在原核启动子的下游，原核 RNA 聚合酶识别该启动子，并带动真核基因在原核细胞中转录。

大肠杆菌表达载体中通常使用调控型的强启动子，如 *lac* 启动子（乳糖启动子），*trp* 启动子（色氨酸启动子），λ 噬菌体 P_L 启动子（λ 噬菌体的左向启动子），*tac* 启动子（乳糖和色氨酸的杂合启动子）等。

（1）*lac* 启动子　即大肠杆菌乳糖操纵子的启动子。乳糖操纵子模型最早是由 Jacob 和 Monod 于 1961 年提出的，它是由启动基因（启动子 P）、操纵基因（O）和编码 3 个与乳糖代谢有关的酶的结构基因所组成。启动子与操纵基因构成调控序列，在调控序列上游不远处还存在一阻遏蛋白基因（*lac*I），其编码的蛋白对该操纵子起调控的作用（图 7-3）。

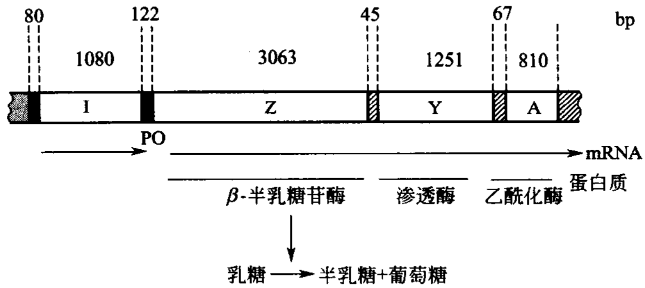


图 7-3　乳糖操纵子结构

此操纵子含有 3 个编码酶蛋白基因 *lac*Z、*lac*Y、*lac*A，它们在乳糖代谢中起不同作用。Z 基因的产物为 β-半乳糖苷酶，它可水解乳糖生成葡萄糖和半乳糖，供细菌利用。Y 基因的产物是通透酶，此酶可使乳糖进入细胞内。A 基因的产物为半乳糖苷乙酰化酶，使半乳糖发生乙酰化。*lac* I 基因为编码阻遏蛋白的基因，而操纵基因（O）是阻遏蛋白结合部位，启动子（P）是转录起始时 RNA 聚合酶结合部位，P 与 O 有部分重叠。

操纵基因（O）可以与四聚体形式的阻遏蛋白结合，从而在空间上妨碍了 RNA 聚合酶的转录。所以，处在阻遏状态下的操纵子就不能产生与乳糖代谢有关的酶。加入诱导物（如乳糖或其衍生物，如 IPTG）后，可与阻遏蛋白形成复合物，而使阻遏蛋白构象改变，不能再与 O 基因结合，RNA 聚合酶可通过此区域，基因即可表达，转录出 mRNA 链，继而翻译出相关的蛋白质。

在启动子上游还存在一个 CAP-cAMP 结合位点。当细胞内缺乏葡萄糖时，会使 cAMP 的浓度上升，与 CAP（分解代谢物活化蛋白）结合形成 CAP-cAMP 复合物，继而结合于乳糖操纵子的 CAP-cAMP 结合位点。CAP-cAMP 复合物是一个正调控因子，对转录起促进作用。但当有大量葡萄糖供应，cAMP 浓度会下降，无法形成 CAP-cAMP 复合物，从而影响了乳糖操纵子的结构基因的表达。

*lac*UV 是一种突变的乳糖启动子，对分解代谢抑制不敏感，即使在没有 CAP-cAMP 的情况下，转录也能照常进行。另外 IPTG 是 β-半乳糖苷酶底物的类似物，有很强的诱导力，它能与阻遏蛋白结合，促进转录。有实验证明，用 *lac* 启动子组建的载体在原核细胞中表达时，IPTG 可提高真核基因的表达水平 50 倍。

（2）*trp* 启动子　即大肠杆菌的色氨酸操纵子的启动子。色氨酸操纵子结构基因排列及其表达如图 7-4。

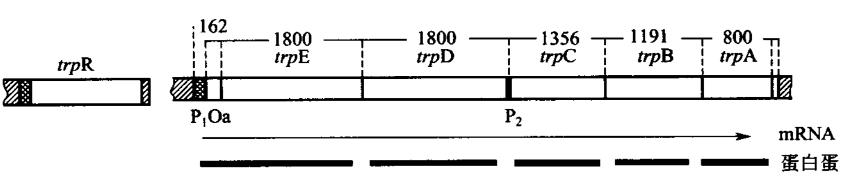


图 7-4　色氨酸操纵子结构

色氨酸操纵子由阻遏物蛋白基因（*trp*R）、启动子（P_1、P_2）、操纵基因（O）、衰减子和结构基因组成。结构基因编码的酶催化分支酸经 5 步反应合成 L-色氨酸。另外，在结构基因 A 之后有两个终止结构 t 和 t′，其中 t′为 ρ 因子所识别。ρ 因子是一个分子量为 200000

的四聚体，能与 RNA 聚合酶结合，帮助其识别终止信号。

启动子 P_1 与操纵基因 O 大部分重叠。P_1 是色氨酸操纵子的主要启动子，它启动结构基因 *trp*E、*trp*D、*trp*C、*trp*B、*trp*A 的转录；另有一个弱启动子 P_2 位于 *trp*D 基因下游，它只控制 3%的 *trp*C、*trp*B、*trp*A 基因的表达，生理作用不大。此操纵子还有一个特点，即调控区不与结构基因 E 直接相连，中间有一段前导肽（L）相隔。L 基因约有 162bp，能编码出 1 个 14 肽。前导肽中有两个色氨酸，这两个色氨酸很重要，因它能与这段转录产物 mRNA 作用形成独特结构，类似转录终止信号，故被命名为衰减子。

trp 启动子受两种调控。阻遏物蛋白基因（*trp*R）平时不断合成阻遏蛋白，只有阻遏蛋白与色氨酸结合，才能作用于操纵基因，阻止转录的进行。所以 *trp* 启动子的调控主要取决于色氨酸的存在于否。当色氨酸缺少时，因上述作用不存在，故操纵子基因启动转录。这种调控类似于 *lac* 启动子的负调控。*trp* 启动子的另一种调控是通过启动子与结构基因间的衰减子进行的。当细胞内色氨酸丰富时，转录到衰减子区域停止；当细胞内色氨酸贫乏，则转录可能通过衰减子区域，进行结构基因的转录。β-吲哚丙烯酸是色氨酸的竞争抑制剂，它能与阻遏蛋白结合，阻止了色氨酸与阻遏蛋白结合，因而能使转录顺利进行。

用于原核表达载体的 *trp* 启动子常常包含启动基因，操纵基因和部分色氨酸 *trp*E 基因，删除了衰减子，则可使转录水平提高 8～10 倍。

（3）*tac* 启动子　亦称 *trp-lac* 启动子，是非常强的启动子。*tac* 启动子是一组由 *lac* 和 *trp* 启动子人工构建的杂合启动子。其中 *tac*1 由 *trp* 启动子的 －35 区加上一个合成的 46bpDNA 片段（包括 Pribnow box 区）和 *lac* 操纵基因构成，而 *tac*12 是由 *trp* 启动子的 －35 区和 *lac* 启动子的 －10 区，加上乳糖操纵子中的操纵基因部分、SD 序列融合而成。*tac* 启动子受乳糖操纵子的阻遏蛋白的负调控，当存在有高水平的 *lac* 阻遏蛋白时，*tac* 启动子受抑制，如加入 IPTG 后，可使 *tac* 启动子去阻遏并诱导其表达。

（4）P_L 和 P_R 启动子　即 λ 噬菌体的左向和右向启动子，它们都是强启动子，比 *lac* 启动子的活性高 8～10 倍。λ 噬菌体基因组的转录是分两期完成的。早期转录中 mRNA 翻译出的蛋白质有两种，即向右转录翻译 Cro 蛋白和向左转录翻译 N 蛋白。两种蛋白质都是作用于操纵基因 O 区的调控蛋白，可加强 P_L 启动子和 P_R 启动子的作用。P_L 和 P_R 启动子受 λ 噬菌体 cI 基因产物的负调控。cI 基因产物是温度敏感蛋白，是 P_L 和 P_R 启动子转录的阻遏蛋白。在较低温度（28～32℃）培养时，cI 基因产物产生抑制作用，在温度升至 42℃时，其被破坏，于是 P_L 和 P_R 启动子开始转录。

P_L 和 P_R 启动子的 Pribnow box 的序列都是 GATAAT，－35 区的序列都是 TGACTA。

在构建带有 P_L 或 P_R 启动子表达载体时，需将 cI 基因组装在表达载体上，或选择溶原化 λ 噬菌体的大肠杆菌作为该表达载体的宿主菌，如 $N99cI^+$ 菌株。

（5）T7 噬菌体启动子　这是一个表达效率很高的启动子，但 T7 噬菌体启动子的转录完全依赖于 T7 噬菌体 RNA 聚合酶，所以在构建带有 T7 噬菌体启动子表达载体时，需要使用能表达 T7 噬菌体 RNA 聚合酶的受体菌，如溶源菌 JM109（DE3）、BL21（DE3）、HMS174(DE3）等。噬菌体 DE3 是 λ 噬菌体的衍生株，一段含有 *lac*I、*lac*UV5 启动子和 T7 RNA 聚合酶基因的 DNA 片段被插入其 *int* 基因中。其调控方式类似于 *lac* 表达系统，可用 IPTG 诱导。通过诱导宿主菌表达的 T7 噬菌体 RNA 聚合酶，来调控 T7 噬菌体启动子的开放和关闭。

（6）其他启动子　除以上几种最常用的启动子外，还有一些其他类型的启动子用于表达

系统的构建，如营养调控型、糖原调控型、pH 调控型、溶氧调控型和生物素调控型等。这些表达系统的特点是通过对菌体发酵和代谢过程条件的控制，实现对目标蛋白表达的调控，是大肠杆菌表达系统走向产业化的发展方向。

2. 核糖体结合位点

核糖体结合位点（ribosome-binding site，RBS）是指紧靠启动子下游、从转录起始位点开始延伸几十个 bp 长度的一段序列，翻译起始密码子 ATG 通常位于它的中心位置。

mRNA 在细菌中的翻译效率严格依赖于是否有核糖体结合位点的存在，这是因为细菌在翻译水平上的调控是不严格的，只有 mRNA 和核糖体的结合，才是蛋白质合成的关键。1974 年 Shine 和 Dalgarno 首先发现，在 mRNA 上有核糖体的结合位点，同时发现在起始密码子 ATG 上游 3～10bp 处存在着一个由 3～9bp 组成的一致性序列（consensus sequence）。这段序列富含嘌呤核苷酸，刚好与 16S rRNA 3′末端的富含嘧啶核苷酸的序列互补，是核糖体 RNA 的识别与结合位点。根据发现者的名字，命名为 Shine-Dalgarno 序列，简称 SD 序列（图 7-5）。

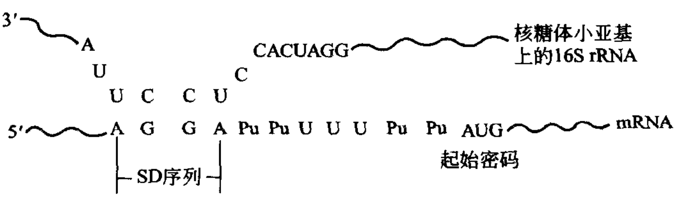


图 7-5　原核生物 mRNA 与核糖体结合的分子机制

SD 序列与起始密码子之间的距离，是影响 mRNA 转译成蛋白的重要因素之一。有人发现当 *lac* 启动子的 SD 序列距 ATG 为 7 个核苷酸时，目的基因的表达最高，而间隔 8 个核苷酸时，表达水平下降 500 倍。这说明 SD 序列与 ATG 的距离将显著地影响基因的表达水平。另外，某些蛋白质与 SD 序列结合也会影响 mRNA 与核糖体的结合，从而影响蛋白质的翻译。

3. 转录终止信号

在一个基因的 3′末端或一个操纵子的 3′末端往往还有一特定的核苷酸序列，它有终止转录的功能，这一 DNA 序列称为转录终止子或简称终止子（terminator）。对 RNA 聚合酶起强终止作用的终止子在结构上有一些共同的特点，有一段富含 A/T 的区域和一段富含 G/C 的区域，G/C 富含区域又具有回文对称结构，这段终止子转录后形成的 RNA 具有茎环结构。并且有与 A/T 富含区对应的一串 U（图 7-6）。转录终止的机制较为复杂，并且结论不统一。在构建表达载体时，为了稳定载体系统，防止克隆的外源基因表达干扰载体的稳定性，一般都在多克隆位点的下游插入一段很强的 *rrn*B 核糖体 RNA 的转录终止子。

（二）大肠杆菌的表达载体

在大肠杆菌中表达外源基因时，由于载体上所带的序列不同，一些载体可用于表达融合型蛋白，一些载体可用于表达非融合型蛋白，还有一些载体可用于表达分泌型蛋白。

在目的基因编码的蛋白质的 N 端连接着由载体的 DNA 序列编码的蛋白质或多肽，这样形成的蛋白质称为融合蛋白。这是由于目的基因在插入到多克隆位点时，在其起始密码 ATG 上游和载体的启动子-SD 序列下游之间的序列中存在着 ATG 序列，可以被细菌的翻译系统识别，从而引起蛋白质的合成，这样在目的蛋白的 N 端就会先合成出一段蛋白质或多

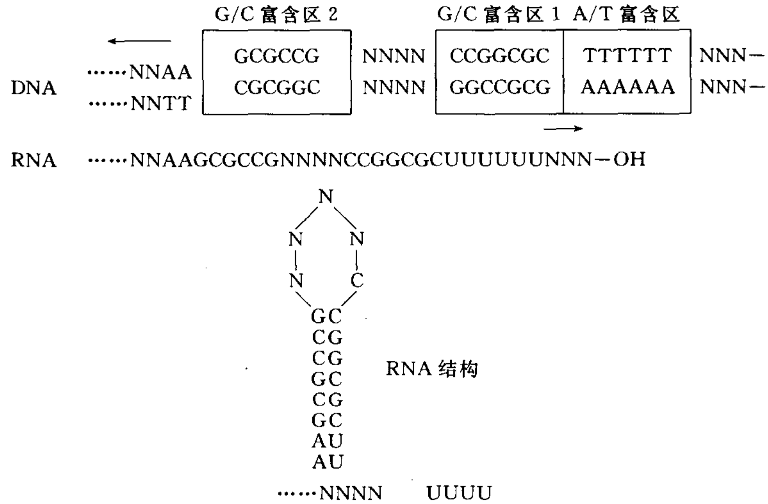


图 7-6　强终止子模式图

肽，从而形成融合蛋白。含原核细胞多肽的融合蛋白是避免被细菌蛋白酶破坏的最好措施。表达融合型蛋白时，为了得到正确的真核蛋白，在插入真核基因时，应非常注意其阅读框架，其阅读框架应与融合的 DNA 片段的阅读框架一致，翻译时才不至于产生移码突变。

不与细菌的任何蛋白或多肽融合在一起的表达蛋白称为非融合蛋白，非融合蛋白的优点在于它具有非常近似于真核生物体内蛋白质的结构，因此表达产物的生物学功能也就更接近于生物体内天然蛋白质。非融合蛋白的最大缺点是容易被细菌蛋白酶破坏。为了在原核细胞中表达出非融合蛋白，可将带有起始密码 ATG 的真核基因插到原核启动子和 SD 序列的下游，组成一个杂合的核糖体结合区，经转录翻译，得到非融合蛋白。

在大肠杆菌中某些蛋白质属于分泌性蛋白质（secretory proteins），可以被分泌到细胞膜外的间质中，这是由于在这些蛋白质的 N 端存在着一段称为信号肽（signal peptide）的多肽。在信号肽的帮助下，蛋白质可以跨过细胞膜而分泌到胞外。因此，将编码信号肽的 DNA 序列插入到载体的启动子与目的基因之间，就可使目的基因编码的蛋白质分泌到胞外，减少了细菌蛋白酶的破坏，同时为表达产物的分离纯化等提供了极大的方便。

1. 融合型蛋白表达载体

(1) pRSET 载体　该载体携带有大肠杆菌复制起始位点，可以在大肠杆菌中复制和扩增。有 *Amp* 抗性基因，可用于转化菌的筛选。有一个强的 T7 噬菌体启动子，由 T7 噬菌体 RNA 聚合酶所识别。在启动子和终止信号之间有 RBS 和多克隆位点（mutiple cloning sites，MCS）。RBS 位点在有的载体上称为 SD 序列。在 RBS 位点之后紧接的是起始密码子（ATG），在 ATG 与多克隆位点之间有一段大肠杆菌基因序列，该序列编码一段多肽。这一序列是为了方便表达产物的纯化。在载体携带的这种序列中，有的含有 6 个组氨酸序列，可以通过镍柱（Ni^{+}）进行亲和层析，纯化表达产物，有的已有相应的单克隆抗体，表达产物可以通过抗体亲和层析柱进行纯化。在这个序列的羧基端，有一个可以被蛋白酶识别的水解位点。在 pRSET 载体中，这一段序列为 31 个氨基酸，其 C 端有一个肠激酶识别位点。可以用相应的酶水解除去大肠杆菌编码的短肽，得到表达产物的天然蛋白（图 7-7）。

为了获得正确编码的表达蛋白，已商品化的 pRSET 载体一般都为一套，一套有 3 个载

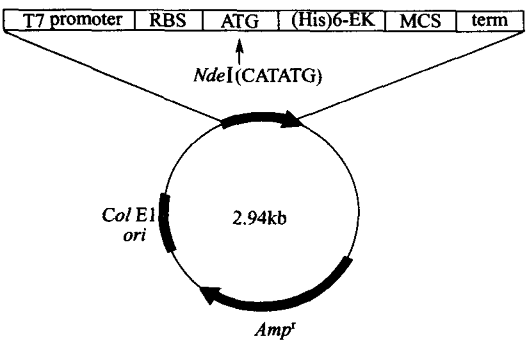


图 7-7　pRSET 载体的物理图谱

体，每个载体除阅读框不同外，其他结构完全一样。如 pRSET 载体分别有 pRSET-A、pRSET-B 和 pRSET-C 可供使用者选择。这样无论外源基因的阅读框架如何，一套载体中总有一个与之匹配。

（2）pGEX 载体　该载体是 Amersham 公司出品的融合蛋白表达载体系统，由 13 种载体（如 pGEX-1λT、pGEX-2T、pGEX-3X、pGEX-4T-1、pGEX-5X-1、pGEX-6P-1）和一种用于纯化表达蛋白的亲和层析介质（glutathione Sepharose 4B）组成。pGEX 载体的组成成分基本上与其他表达载体相似，含有启动子（*tac*）及 *lac* 操纵基因、SD 序列、*lac*I 阻遏蛋白基因等。这类载体与其他表达载体不同之处是 SD 序列下游就是谷胱甘肽-S-转移酶基因，而克隆的外源基因则与谷胱甘肽-S-转移酶基因相连。当进行基因表达时，表达产物为谷胱甘肽-S-转移酶和目的基因产物的融合体。这个载体系统具有如下优点：①可诱导高效表达；②载体内含有 *lac*I 阻遏蛋白基因；③表达的融合蛋白质纯化方便；④使用凝血酶和Ⅹa 因子就可从表达的融合蛋白中切下所需要的蛋白质和多肽；⑤用 *Eco*RⅠ从 λgt 11 载体中分离的基因可直接插入 pGEX-1λT 中。

2. 非融合型蛋白表达载体

用于表达非融合型蛋白的表达载体关键在于两个方面。①载体的 SD 序列下游的适当位置必须构建有起始密码子（ATG）。有些载体在 SD 序列的 3′端下游构建了一个 *Nde*Ⅰ（CATATG）或 *Nco*Ⅰ（CCATGG）位点。起始密码子 ATG 是包含在 *Nde*Ⅰ或 *Nco*Ⅰ位点中的。如果目的基因的 5′端为 *Nde*Ⅰ或 *Nco*Ⅰ位点，当其与载体连接时，不仅便于插入，而且也引入了起始密码子。②ATG 与载体上的 SD 序列要有合适的距离，因为载体上的 SD 序列与 ATG 之间的距离是影响蛋白质翻译效率的一个非常重要的因素。凡具备这些特点的载体都可作为表达非融合型蛋白的载体。

（1）pRSET 载体　该载体是一个两用载体。在该载体 SD 序列 3′端的下游构建了一个 *Nde*Ⅰ位点，因此该质粒即可作为表达融合型蛋白的载体，也可作为表达非融合型蛋白的载体（图 7-7）。

（2）pKK223-3 载体　该载体具有一个强的 *tac* 启动子。这个启动子是由 *trp* 启动子的 −35 区和 *lac*UV 5 启动子的 −10 区、操纵基因及 SD 序列组成，可以用 IPTG 诱导表达，如图 7-8 所示，紧接 *tac* 启动子的是一个取自 pUC8 的多克隆位点，使之很容易把目的基因定位在启动子和 SD 序列后；在多克隆位点下游的一段 DNA 顺序中，还包含一个很强的 *rrn*B 核糖体 RNA 的转录终止子，目的是为了稳定载体系统。因为上游强的 *tac* 启动子控制的转录，必须由强终止子抑制，才不至于干扰和载体本身稳定性有关的基因表达，载体其余部分是由 pBR322 组成的。pKK223-3 载体的宿主应具有 *lac* I^+ 基因型，如 JM105。

3. 分泌型蛋白表达载体

pIN-ompA 载体是以 pBR322 为基础构建的，它含有能在大肠杆菌中复制的复制起始位点、用于筛选的抗性基因、带有 Ipp（脂蛋白基因）启动子，这是大肠杆菌中最强的启动子之一，在启动子的下游装有 *lac*UV-5 的启动子及其操纵基因，另外还将 *lac* 阻遏蛋白的基因

（*lac* I）也克隆在载体上，使目的基因的表达成为可调节表达（图 7-9）。

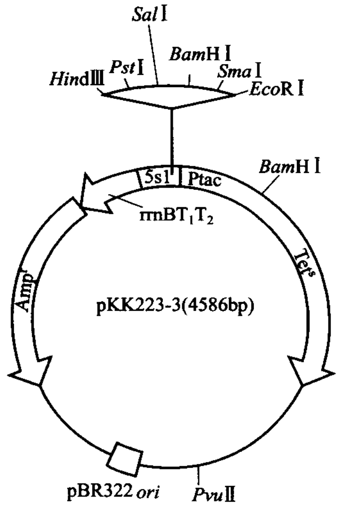


图 7-8 pKK223-3 载体的物理图谱

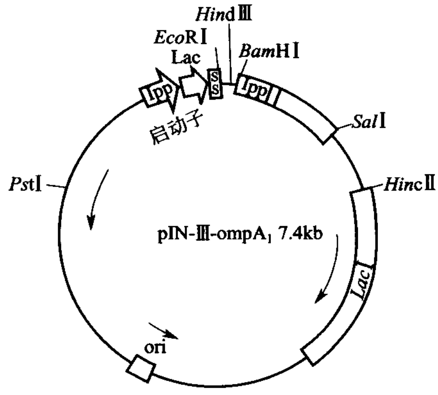


图 7-9 pIN-ompA 载体的物理图谱

在转录控制顺序的下游再装上人工合成的高效翻译起始顺序（SD 序列及 ATG）。作为分泌克隆表达载体中关键的编码信号肽的顺序，是取自于大肠杆菌中分泌蛋白的基因，ompA（外膜蛋白基因）。在信号肽编码顺序下紧接着的是一段人工合成的多克隆位点片段，其中包含 3 个单酶切位点：*Eco*RⅠ、*Hin*dⅢ和 *Bam*HⅠ。为了适应带不同阅读框的目的基因片段与信号肽密码子阅读框完全匹配，避免移码突变，已构建了一套载体，分别为 pIN-ompA$_1$、pIN-ompA$_2$ 和 pIN-ompA$_3$。用这种分泌型载体可以将表达产物分泌到大肠杆菌的细胞间质中，信号肽可被信号肽酶所切割。

（三）外源基因在大肠杆菌中表达

在大肠杆菌中表达外源基因，主要考虑以下一些基本要素。

1. 载体

所用载体必须是大肠杆菌表达载体，具有大肠杆菌的复制起始位点和抗性基因，有一个强的原核启动子（如 P_L、*tac*、T7 等）及其两侧的调控序列，有 SD 序列，而 SD 序列与起始密码子 ATG 之间要有合适的距离；载体的启动子与目的基因有正确的阅读框以及载体的 3′端必须构建有转录终止信号。

2. 目的基因

插入大肠杆菌表达载体的目的基因可以是真核基因，也可以是原核基因。如果目的基因来自真核细胞，必须是 cDNA，因为大肠杆菌没有剪切内含子的功能。从真核基因转录的 mRNA 缺乏结合细菌核糖体的 SD 序列，因此，cDNA 的起始密码子（ATG）上游部分（5′端非编码区）是无用的，必须除去。对于一些分泌性蛋白，还应除去信号肽部分。

3. 重组 DNA 表达载体的构建和表达

对于 pRSET 两用载体来说，目的基因与载体有两种不同的连接方式，可构建成两种不同的重组 DNA 表达载体，产生两种不同的蛋白，即融合型蛋白和非融合型蛋白。

（1）产生非融合型蛋白时目的基因与载体的连接方式 利用 pRSET 载体的 SD 序列 3′端下游的适当位置构建的 *Nde*Ⅰ（CATATG）位点和 MCS 中的一个酶切位点将载体切开，

去掉载体上携带的一段大肠杆菌基因序列，将5′端带有*Nde* Ⅰ酶切位点和3′端带有另一种酶切位点的目的基因与载体相连，构建成重组DNA表达载体。作为目的基因的cDNA，自身必须含有起始密码子（ATG）和终止密码子（TGA）。*Nde* Ⅰ位点中的ATG即可作为起始密码子，又是一个酶切位点，便于DNA片段的连接和切割。这种构建方式适合在大肠杆菌中表达非融合蛋白。

非融合蛋白的优点在于它具有非常近似于真核生物体内蛋白质的结构，因此表达产物的生物学功能也非常接近于生物体内天然蛋白质。

（2）产生融合型蛋白时目的基因和载体的连接方式　利用pRSET载体MCS中的酶切位点将载体切开，将带有相同酶切位点的目的基因片段与载体相连。目的基因的5′端正好与载体上携带的大肠杆菌基因序列的3′端相连。这种构建方式适合在大肠杆菌中表达融合型蛋白。

在表达融合蛋白时，为得到正确编码的表达蛋白，在插入外源基因时，其阅读框应与载体上的原核DNA片段的阅读框一致，这样，插入的外源基因在翻译时才不致产生移码突变。

表达融合蛋白有以下几个优点：①融合蛋白较稳定，不易被细菌蛋白酶水解；②如果大肠杆菌的结构基因是一段信号肽，可产生分泌型产物；③可利用针对原核部分的单抗进行亲和层析，便于纯化；④原核蛋白部分可用蛋白酶切掉，释放出天然的真核蛋白质。

（3）产生分泌型蛋白时目的基因和载体的连接方式　产生分泌型蛋白时目的基因和载体的连接方式与产生融合型蛋白时目的基因和载体的连接方式一样。如选用pIN-ompA载体，在该载体的MCS中选用两个酶切位点将载体切开，然后与带有与其相同酶切位点的目的基因相连，目的基因的5′端正好与载体上携带的信号肽序列的3′端相匹配。这种构建方式适合在大肠杆菌中表达分泌型蛋白。同样要注意在插入外源基因时，其阅读框应与信号肽的阅读框一致。

蛋白质能够在大肠杆菌中进行分泌，至少要具备3个要素：①有一段信号肽；②在成熟蛋白质内有适当的与分泌相关的氨基酸序列；③细胞内有相应的转运机制。

4. 受体菌株和诱导条件

除了表达载体以外，受体菌株也是构成原核表达系统的重要组成部分之一。因此，合理地选择使用受体菌株对于外源基因是否能够表达和表达效率的高低起到决定性的作用。选择使用受体菌株的原则主要是根据启动子受何种阻抑物的调节来决定的。诱导条件也要根据启动子类型和特定的蛋白质而定。为了提高外源基因的表达水平，常用的方法是将受体菌的生长和外源基因的表达分开。

带有*tac*启动子的载体，受控于*lac*阻抑物，因此应选择能表达*lac*阻抑物的*lac*I^q大肠杆菌菌株作为受体菌株。将该菌株置37℃培养时，能高效表达*lac*阻抑物，使*tac*启动子受到抑制，外源基因不能表达。此时，受体菌得到大量生长。当加入IPTG后，*lac*阻抑物不能与操纵基因结合，外源基因大量转录并高效表达。带有*tac*启动子的载体能选用的受体菌有JM109、XL-1-blue、RB791等。

带有P_L启动子的载体，要求宿主菌能表达cIts857阻抑物。将含有P_L启动子载体的M5219菌株置30℃培养时，该受体菌可以高水平地表达cIts857阻抑物，使载体上的P_L启动子受到抑制，外源基因不能表达，此时，受体菌不断地增殖，同时载体也得到大量地扩增。然后，再将该受体菌置42℃继续培养。在42℃培养条件下，cIts857阻抑物不能表达，

P_L 启动子解除阻遏，外源基因得以高水平表达。

带有 T7 启动子的载体，则需要宿主菌带有 T7 噬菌体 RNA 聚合酶基因，该基因受 *lac*UV5 启动子的调控。受体菌在 37℃进行培养，细菌大量扩增。加入 IPTG 后可诱导 T7 噬菌体 RNA 聚合酶基因的表达，表达的 T7 噬菌体 RNA 聚合酶可识别载体上的 T7 启动子，使其下游基因表达。

另外，对于一个特定的蛋白质来说，并不是所有的菌株都能获得相同的表达效率，有时需要试几种菌株，以选择最好的一种。

（四）提高外源基因表达水平的措施

采用基因工程技术，将外源基因与大肠杆菌表达载体重组连接后，导入大肠杆菌，目的基因可在大肠杆菌中高效表达。由于大肠杆菌增殖速度很快、培养条件简单、能够制备出大量的表达产物，因此除有些必须要进行翻译后加工才具有生物学活性的真核蛋白外，大多数蛋白质都可以利用原核表达系统进行表达。如前所述，外源基因在大肠杆菌中表达应考虑选择表达载体的类型、阅读框、所用启动子的强弱、目的基因的制备、载体与目的基因的连接以及宿主菌株和诱导条件等因素。只要考虑到这些因素，一般都可以使外源基因在大肠杆菌中有所表达。但要使外源基因在原核细胞中获得高效表达，还应考虑采取以下措施。

1. 提高翻译水平

（1）增加 mRNA 的稳定性　外源基因 mRNA 的半衰期很短，容易被降解，以致影响外源基因的表达水平。可考虑在外源基因下游插入“重复性基因外回文”（repetitive extragenic palindromic，REP）顺序。该顺序能防止 3′→5′外切酶的攻击，具有稳定 mRNA、提高表达水平的作用。

（2）调整 SD 序列与 ATG 之间的距离　mRNA 链中 SD 序列和起始密码子 ATG 之间的距离是影响基因表达水平的一个重要因素。距离过长、过短都不利于基因表达。对于不同的基因以及不同的启动子，最适距离是不一样的，一般为 5～9bp。增加一个或减少一个 bp，表达效率可以降低几倍乃至几百倍。SD 序列与 ATG 之间的距离可采用基因工程的方法进行调整。

（3）用点突变的方法改变某些碱基　有资料证明，起始密码子下游的几个密码子如果采用不同的核苷酸组成，可使基因的表达效率相差 15～20 倍。这主要是改变了翻译的起始和 mRNA 的二级结构。利用密码子的简并性，采用化学合成的方法，合成一组合适的密码子，可以提高翻译效率。

2. 使细菌的生长与外源基因的表达分开

外源基因在细菌中高效表达，必然影响宿主的生长和代谢；而细胞代谢的损伤，又必然影响外源基因的表达。合理地调节好宿主细胞的代谢负荷与外源基因高效表达的关系，是提高外源基因表达水平不可缺少的一个环节。

将宿主菌的生长和外源基因的表达分开成为两个阶段，是减轻宿主细胞的代谢负荷最为常用的一个方法。一般采用温度诱导或药物诱导。在特定条件下，培养含重组 DNA 表达载体的受体菌，由于细菌的生长大约 20min 繁殖一代，此时受体菌不断增殖，同时也使转化到受体菌中的重组 DNA 载体得到大量扩增。当细菌生长到所需浓度时，在培养液中加入诱导剂或改变培养温度，使细菌生长速度减慢，以外源基因的表达为主。如前所述，含 *tac* 启动子、P_L 噬菌体启动子和 T7 噬菌体启动子的重组 DNA 载体分别导入大肠杆菌进行表达时，都是采取将细菌的生长与外源基因表达分开的方法；这样可大大减轻宿主细胞代谢负

荷，提高外源基因的表达水平。

减轻宿主细胞代谢负荷的另一个措施，是将宿主菌的生长和表达质粒的复制分开。当宿主菌迅速生长时，抑制质粒的复制；当宿主菌生物量积累到一定水平后，再诱导细胞中质粒DNA 的复制，增加质粒的拷贝数，拷贝数的增加必将伴随外源基因表达水平的提高。

3. 提高表达蛋白的稳定性

在大肠杆菌中表达的外源蛋白常常被菌内蛋白酶降解，导致外源基因的表达水平大大降低。为了提高表达蛋白的稳定性，防止细菌蛋白酶的降解可采取以下措施。

（1）表达融合蛋白　融合蛋白较稳定，不易被细菌蛋白酶水解。

（2）采用某种突变株　可采用大肠杆菌蛋白酶缺陷型菌株作受体菌（如 *lon* 缺陷型菌株），可使大肠杆菌蛋白酶合成受阻，从而使蛋白质得到保护，不被降解。也可以将编码细菌蛋白酶抑制剂的基因（如 T4 噬菌体的 *Pin* 基因）克隆到质粒中，将此质粒转化到受体菌中，其产物可使细菌的蛋白酶受到抑制，使表达蛋白受到保护。

（3）表达分泌蛋白　表达分泌蛋白的载体，在起始密码子后有一段编码信号肽的序列，所产生的融合蛋白 N 端的原核序列即为信号肽。此表达蛋白可从胞质跨过内膜进入周间质，防止了宿主菌对表达蛋白的降解，同时也减轻了宿主菌代谢负荷，并且使表达产物恢复天然构象。

（五）包含体

在原核细胞表达外源基因，尤其是以大肠杆菌为宿主菌高效表达外源基因时，表达蛋白常常在细胞质内聚集，形成包含体（inclusion body）。包含体的形成有利于防止蛋白酶对表达蛋白的降解，并且非常有利于分离表达产物。但包含体形成后，表达蛋白不具有物理活性，因此必须溶解包含体并对表达蛋白进行复性。包含体形成后另一个不利方面是，由于表达产物形成包含体，负责水解起始密码子编码的甲硫氨酸的水解酶，不能对所有的表达蛋白质都起作用，这样就可能产生 N 末端带有甲硫氨酸的目的蛋白质的衍生物，而非生物体内的天然蛋白，这可能会对某些蛋白质的性质产生影响。

包含体是具有膜结构的非结晶性蛋白质聚集体。在偏振光显微镜或电子显微镜下可发现包含体与细胞质的明显区别。不仅带有重组高效表达质粒的大肠杆菌可形成包含体，而且有变异蛋白产生或菌体本身蛋白质异常高表达的大肠杆菌也可形成包含体。包含体的形成原因和形成过程还不十分清楚。有人认为包含体的形成不仅与表达蛋白的生成速率高、无足够时间使肽链折叠及表达蛋白的高浓度有关，而且还与培养宿主菌的温度、pH、某种金属离子不足等造成细胞内环境变化等因素有关。

常规方法使菌体破碎（如超声波、匀浆等）后，离心就可得到包含体。密度梯度离心后则可得到高纯度的包含体。包含体通常不溶于水，加入强蛋白质变性剂后方可溶解。根据包含体中蛋白质种类的不同，通常用 6～8mol/L 盐酸胍或 9～10mol/L 尿素溶解包含体。有时不用蛋白质变性剂，而采用 pH2.0～4.5 的酸性溶液也可使包含体溶解。

表达蛋白的肽链中含有半胱氨酸时，在包含体内可形成二硫键。菌体破碎后，由于空气的氧化作用，表达蛋白分子间或与杂蛋白质也可形成二硫键。这种情况下，用变性剂溶解包含体时，再加入二硫苏糖醇（DTT）、β-硫基乙醇等还原剂可使表达蛋白质完全还原。使形成包含体的表达蛋白恢复其活性的过程称为包含体蛋白质复性，其基本原理是随着变性剂浓度的降低，表达蛋白质恢复其天然构象。降低变性剂浓度的方法有多种，常用的有稀释法、透析法、凝胶过滤及各种层析方法等。最简单而又常用的方法是稀释法或透析法。

总之，为了提高蛋白质复性效率，在溶解包含体时应：①尽量提高包含体的纯度；②尽量使包含体充分溶解，并且使形成包含体的蛋白质完全还原；③必要时加入还原剂。在具体复性实验过程中应：①分阶段减低变性剂浓度，并使变性剂保持在适当浓度；②使用氧化还原缓冲液；③使复性的蛋白浓度、pH、温度及盐的种类、离子强度等最适。当然，如果能防止表达蛋白形成包含体，是保存蛋白质功能最好的方法。目前，这方面的研究已正在进行中，基本的设想是与目的基因同时表达分子伴侣蛋白。

二、芽孢杆菌表达系统

芽孢杆菌为革兰阳性细菌，细胞壁不含内毒素，能分泌大量蛋白到体外，为一些重要工业酶制剂的生产菌种，能形成芽孢，而子囊中只有一个芽孢。除个别种类如炭疽芽孢杆菌（*Bacillus anthracis*）和蜡样芽孢杆菌（*Bacillus cereus*）外，其他芽孢杆菌对人畜无毒。芽孢杆菌表达系统是在 20 世纪 70 年代从枯草芽孢杆菌，又称枯草杆菌（*Bacillus subtilis*）开始，逐步扩展至其他种。迄今，已在枯草杆菌及其近缘种中克隆和表达了大量原核和真核基因，其中有的已应用于工业生产，取得了不少成绩。芽孢杆菌表达系统有许多特点，同时也暴露出该系统存在蛋白酶和质粒不稳定等问题。

（一）芽孢杆菌基因表达的特点

1. 形成芽孢

芽孢杆菌在对数生长期后就进入芽孢形成期，一个细胞分化成两个细胞，其中一个形成芽孢，另一个不形成芽孢。不同的芽孢形成阶段，表达的基因不同。当然，对数生长期表达的基因与芽孢形成期表达的基因也不相同。为了研究芽孢形成的分子机理以及便于在不同阶段尽量高地表达克隆基因，在枯草杆菌中分离了于各个芽孢形成期停止形成芽孢的各种突变体。这些突变体都可从美国俄亥俄州立大学芽孢杆菌遗传保存中心（BGSC）获得。

2. 多 σ 因子

原核生物基因的转录主要由 RNA 聚合酶来完成。该酶由 5 个亚基组成：2 个 α 亚基、1 个 β 亚基、1 个 β 亚基和 1 个 σ 因子。5 个亚基组成 RNA 聚合酶全酶，除 σ 因子外其余 4 个亚基构成核心酶。σ 因子的主要功能是帮助核心酶识别特定的启动子而结合到转录的起始部位。转录开始后 σ 因子就不起作用了。枯草杆菌 σ 因子已发现有 σ^A（σ^{43}，σ^{55}）、σ^B（σ^{37}）、σ^D（σ^{28}）、σ^E、σ^H（σ^{29}）、σ^{gp28}和 $\sigma^{gp33\sim34}$；而大肠杆菌只有 σ^{70}、σ^{32}。这种多 σ 因子与芽孢形成有关，也与重叠启动子、非重叠双启动子以及转录级联有关。芽孢杆菌中存在着双启动子基因或操纵子，有的是重叠的，有的是非重叠。这些双启动子可加强基因的表达。苏云金芽孢杆菌杀虫晶体蛋白基因、地衣芽孢杆菌细胞壁表面蛋白基因都有双启动子。

3. 16S RNA 3′端序列

在翻译时，16S rRNA 3′端与核糖体识别 mRNA 并起始翻译的过程有关。枯草杆菌 16S rRNA 的 3′端序列与大肠杆菌的是不相同的。

4. 基因表达具有革兰染色分类

大肠杆菌为革兰阴性菌，芽孢杆菌为革兰阳性菌。作为革兰阴性菌的大肠杆菌，其基因除个别的以外，一般不能直接在属于革兰阳性菌的芽孢杆菌中表达。而芽孢杆菌的基因可以在大肠杆菌中表达。革兰阳性菌中，分类相差很远的菌，如芽孢杆菌与金黄色葡萄球菌，质粒可共用，基因也可相互表达。有实验表明 S1 蛋白对枯草杆菌中蛋白合成重要性。枯草杆菌 30S 核糖体亚基缺少 S1 蛋白。这种翻译与革兰染色间的关系具有普遍性。例如，革兰阴

性菌铜绿假单胞杆菌和棕色固氮菌也具有S1蛋白而无此翻译特异性；而一些革兰阳性菌也都没有S1蛋白则有此翻译特异性。用大肠杆菌核糖体30S、50S亚基与枯草杆菌的30S、50S亚基进行4种组合，分别翻译革兰阳性mRNA（φ29）和革兰阴性的mRNA（MS2）。当翻译φ29时无论那种亚基组合都可进行翻译（甲硫氨酸掺入），而翻译MS2 mRNA时，只有30S亚基来自大肠杆菌才能有甲硫氨酸掺入；枯草杆菌30S亚基无论与大肠杆菌还是枯草杆菌的50S亚基进行组合，均未观察到甲硫氨酸掺入。

5. 大量分泌蛋白

由于芽孢杆菌和大肠杆菌细胞膜的结构有所不同，而导致它们在分泌蛋白方面的差异。大肠杆菌分泌蛋白往往分泌到两层细胞膜之间，而芽孢杆菌由于只有一层细胞膜，蛋白跨膜之后就分泌到培养基之中，因而可以大量分泌，达到30g/L。蛋白跨膜分泌是一个复杂的分子过程。枯草杆菌有将近300个N端具有信号肽的跨膜蛋白。不过用来表达和分泌克隆基因的信号肽主要是来自蛋白酶、淀粉酶和细胞壁表层蛋白等。

6. 单基因受多调控基因调控

芽孢杆菌的基因，往往受多个调控基因的调控。*apr*E（芽孢杆菌碱性蛋白酶E基因）在其上游调控区段就有8个调控基因的结合部位。这些调控基因以及另有一些还未确定结合位点的调控基因对碱性蛋白酶的表达起着正或负调控作用。

7. 一个调控基因往往影响多个基因的表达

DegQ（HY）、DegT可同时促进蔗糖-6-果糖基转移酶、碱性蛋白酶、中性蛋白酶、α-淀粉酶、胞内丝氨酸蛋白酶、木聚糖酶、纤维素酶高表达。

（二）芽孢杆菌表达系统

芽孢杆菌表达系统具有以下特点：①枯草芽孢杆菌是非致病的土壤微生物，严格生长在有氧条件下，不像大肠杆菌那样具有热源脂多糖；②枯草杆菌遗传学现在相当先进，很多噬菌体和质粒适合于用作克隆载体；③芽孢杆菌可大量产生几个商品酶，如α-淀粉酶、蛋白酶以及苏云金杆菌的杀虫晶体蛋白等，因此发酵技术是相当发达的；④作为革兰阳性成员的枯草杆菌，具有单层细胞膜组成较简单的细胞外壳，当分泌蛋白跨过细胞膜后，就被加工和直接释放到培养基中，这使得回收和纯化分泌蛋白较为简单，很多商业酶是芽孢杆菌产生的胞外蛋白。

1. 宿主菌株

（1）枯草杆菌宿主菌株　将携有外源基因载体导入宿主细菌中是细菌基因工程表达系统的必要条件。由于大肠杆菌的氯化钙转化方法对枯草杆菌无效，所以枯草杆菌基因工程表达系统中的宿主菌株用得最广泛的就是在DNA重组技术发明之前就可进行感受态转化的168菌株及其突变体。这些突变体包括各式各样的营养要求、芽孢形成和萌发、蛋白酶缺失、重组缺陷、限制/修饰系统缺陷、转座子插入等方面的突变体。其中大多数可从BGSC获得。目前该中心保存的突变体数量达到900个以上。

1964年，汤懋竑等鉴定了从全国各地收集的大量枯草杆菌菌株的感受态转化特性，其中江西农学院从江西土壤中分离的Ki-2株为可转化菌株。从Ki-2菌株获得的突变体中，Ki-2-132和Ki-2-148为高频转化株，并建立了宿主/载体表达系统。从转化方法、最高转化频率、质粒分离性的稳定性、与大肠杆菌原生质体融合等几方面比较，Ki-2菌株比168菌株均具有优势。

随着非感受态转化方法（如原生质体转化和电转化等方法）的建立，再加上转导、转染

和接合转移，枯草杆菌宿主菌株就相应地扩大了范围，在没有发现感受态转化菌株中也进行了异源基因表达。但文献中绝大多数宿主菌株仍为可感受态转化菌株。

（2）其他宿主菌株　除枯草杆菌以外，已报道用作宿主表达克隆基因的还有：嗜碱芽孢杆菌、淀粉芽孢杆菌、短芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、短小芽孢杆菌、球形芽孢杆菌、嗜热脂肪芽孢杆菌、苏云金芽孢杆菌、耐碱的芽孢杆菌等。

2. 载体

主要有自主复制质粒、整合质粒和噬菌体三种。

（1）自主复制质粒　从芽孢杆菌中分离的自主复制质粒，除极少数以外（例如 pBC16），均为无抗性标志的隐秘质粒。带有抗性标志的自主复制质粒主要来自其他革兰阳性细菌，特别是来自金黄色葡萄球菌的质粒。其中使用广泛的有 pUB110、pC194、pE194 等。在这些质粒的基础上构建了双标记质粒、芽孢杆菌/大肠杆菌穿梭质粒、表达质粒、整合质粒、探针质粒等。来自金黄色葡萄球菌等革兰阳性菌的这些小质粒进行滚环式复制，复制时产生大量单链 DNA（ssDNA）。由于它的存在，使质粒容易产生丢失而形成结构性不稳定。这种丢失似乎是随机的；当然，在选择压力下有些部位就不随机了。将外源基因整合到宿主染色体中，是目前解决这种不稳定性的有效途径。一些大的革兰阳性质粒如 pAMB1 等是 θ 型复制。

（2）整合质粒　采用整合质粒将克隆基因整合到宿主染色体，是克服芽孢杆菌质粒不稳定性的一个有效途径。这种质粒的基本结构是在大肠杆菌质粒的基础上增加一个芽孢杆菌的抗性标志，以及待整合的目的基因。它在大肠杆菌中进行基因克隆或亚克隆操作。整合质粒导入芽孢杆菌后，由于它没有芽孢杆菌质粒的复制起点而不能自主复制，只有插入到宿主染色体后，随着细胞复制而复制。在含有整合质粒中芽孢杆菌抗性标志的抗生素的培养基上，就很容易挑出这种整合体。整合质粒也存在带有转座子（如 *Tn*917、*Tn*10）或芽孢杆菌温度敏感质粒的其他类型。

（3）噬菌体　不少噬菌体都可用作载体，如 Φ105 噬菌体、SPβ 噬菌体以及其他噬菌体。其中，Φ105 噬菌体应用较多，它是一个温和噬菌体，基因组约为 39.2kb，从中发展了不少载体。

SPβ 为一原噬菌体，基因组约 120kb，枯草杆菌 168 菌株及其衍生株都带有此原噬菌体。枯草杆菌 Ki-2-132 中也发现缺陷的原噬菌体。

（三）存在的问题

在 20 世纪 70 年代末与 80 年代，国内外不少著名的大学、研究所与公司怀着对枯草杆菌工程表达系统的希望，在完善该系统、表达真核基因方面进行了大量的工作。经过十多年的努力，发现枯草杆菌存在着蛋白酶和质粒不稳定性两大难题。枯草杆菌中应用最广泛的宿主菌株——168 菌株，不但蛋白酶分泌量大而且种类多。用基因失活突变的方法失活了染色体上中性蛋白酶、碱性蛋白酶、胞内蛋白酶、金属蛋白酶、芽孢杆菌肽酶 F 和中性蛋白酶 B 6 个蛋白酶基因后，仍有 0.32% 的野生型蛋白酶表达。有人曾用合成的蛋白酶抑制剂来提高胰岛素原在枯草杆菌中的表达量，但这一办法并未得到广泛的采用。枯草杆菌表达真核基因总的是不理想，并未达到人们原来的期望，主要原因是胞外蛋白酶降解真核基因产物。因此，其他缺失蛋白酶基因芽孢杆菌以及本来就不产蛋白酶的芽孢杆菌就成为高表达真核基因的宿主。目前，芽孢杆菌表达真核基因至少有 2 个例子，其表达量均超过大肠杆菌和酵母表达系统的表达量，一是地衣芽孢杆菌表达 IL-3，宿主是经遗传修饰的工业用菌，其表达量

为大肠杆菌和酵母表达的10倍，为哺乳动物细胞的100倍，而且分泌到胞外；二是短芽孢杆菌表达人表皮细胞生长因子（hEGF），表达量大大超过大肠杆菌和酵母表达该产物的量。但后一种表达系统被证实存在不稳定性。综上所述，采用完全没有胞外蛋白酶而可稳定表达的芽孢杆菌基因工程系统来表达真核基因，可能是这方面工作的一个方向。

用芽孢杆菌表达了很多酶，取得了很大进展，有的已商业化，产生了很大的社会和经济效益。现已商业化的主要产品还是蛋白酶和淀粉酶。这主要是芽孢杆菌原来就是这方面的生产菌。基因工程给它以量和质的提高。目前这方面存在的问题主要是质粒的稳定性，为绕过这一问题，主要的办法是将克隆基因整合到染色体。估计今后可能在以下几个方面会有进展：①扩大芽孢杆菌工程菌表达和生产酶的种类；②提高质粒的稳定性，包括分离性和结构性的稳定性；③利用多个调控基因和诱导等方法提高表达量；④将已有的工业用菌株改建成为工程菌等。

三、链霉菌表达系统

链霉菌是一类需氧、丝状的革兰阳性细菌，广泛存在于土壤中，可产生约100000种已知生理活性的物质，包括抗细菌、抗真菌、抗肿瘤、抗病毒、免疫调节、酶抑制剂等。所以，链霉菌常常被看作是制造各种生物活性产物的丰富资源。链霉菌还产生多种蛋白酶，以分解不同的复杂碳源、氮源和磷源。链霉菌从形态分化到次级代谢的产生以及蛋白酶的分泌均受着高度的调控，包括在启动子和RNA聚合酶水平、营养代谢水平、感受器/反应器双元件水平、ppGpp系统水平、调控tRNA水平以及被称为细菌激素的丁酸内酯类化合物系统等的调节。链霉菌产生次级代谢产物的多样性、对蛋白外泌的能力以及复杂的相互偶联的多层次调控机制的特征，作为外源基因的表达系统受到了广泛的关注。目前对链霉菌的生物学及作为基因克隆的宿主和载体系统、基因表达和调控以及蛋白的分泌等方面的研究和应用均取得一定的进展。

（一）链霉菌的生物学特征

链霉菌大多数来自于土壤，是能形成孢子的革兰阳性细菌，在固体培养基上有复杂的形态和生理生命周期。它们通常以无中隔的分支菌丝方式生长，当生长到一定阶段营养耗竭时通常能激活形成孢子的途径，导致气生菌丝和孢子的产生。

链霉菌的基因组约为大肠杆菌的两倍，与细菌不同的是它们的染色体是线性的，约有8Mb，在中间含有复制起始点和共价结合的末端蛋白，其末端有数十个到数百个碱基重复序列。3′端回文序列能形成发夹结构。链霉菌中也发现有线性质粒，其复制是从中间开始，向着两端进行。链霉菌属基因的不稳定性归因于染色体末端的不稳定性，长达2Mb不稳定序列经常有自发的缺失或扩增，有时这种自发的缺失导致染色体的环化，但对生长没有严重影响。

链霉菌可作为宿主菌、利用不同的载体和启动子使目的基因表达，产生胞外蛋白。链霉菌还能产生许多胞外酶，但是其分泌蛋白酶的量不很多，因此对外源蛋白的分泌表达干扰性不太大。

大多数链霉菌的生长阶段呈现两相动力学特征：在菌丝生长期主要是积累菌体量，消耗容易被利用的碳、氮源，进入生长稳定期早期时，次级代谢产物的产生受到全局的调节（如解除碳源、氮源或磷酸盐的阻遏），代谢途径专一的调节基因得到表达或小分子调节因子的基因转录得到活化，使链霉菌进入发育分化阶段，开始产生孢子及次级代谢产物。

(二) 链霉菌基因表达的特点

1. RNA 聚合酶基因多样性

天蓝色链霉菌中 RNA 聚合酶基因多样性是由不同的 σ 因子决定的。现已证实，天蓝色链霉菌中至少有 7 种不同的 σ 因子，如 σ^{32}、σ^{49}、σ^{WhiG}、σ^{F} 等，负责调控天蓝色链霉菌生活周期中不同基因的表达，包括孢子形成和次级代谢产物合成。

含有 σ^{32} 因子的 RNA 聚合酶全酶识别 *veg* 启动子与大肠杆菌启动子有保守性，含有 σ^{49} 因子的 RNA 聚合酶全酶识别发育阶段的启动子 *ctc* 和 *endo*，与大肠杆菌启动子没有保守性，而它们自身有一定保守性。

σ^{WhiG} 在产生孢子方面起着重要作用，而 σ^{F} 控制链霉菌发育的最后阶段——孢子成熟阶段，σ^{F} 的表达直接或间接地依赖于 σ^{WhiG}。

2. 调节基因

在链霉菌次级代谢途径中已知有两种类型调节基因：途径专一调节基因（pathway specific regulator）和全局调节基因（global regulator）。前者是对专一的合成途径有调节作用的基因，如 *red*D 调节十一烷基灵红菌素的合成途径，*str*R 调节链霉素的合成途径，*cca*R 调节头霉素和克拉维酸的合成途径，*srm*R 调节螺旋霉素的合成途径等。以上这些途径专一调节基因的产物均属于链霉菌调节蛋白新家族（SARP）。这一类蛋白在其 N 端有与 OmpR 相似的 DNA 结合区。SARP 在许多链霉菌中都存在，对调节许多抗生素的产生起着主要作用。这些正调控基因的调控有时序性，即在菌丝生长稳定期及抗生素开始产生时启动转录。

全局调节基因是指对产生菌中所有抗生素生物合成起调控作用的基因，如 *abs*A 和 *abs*B 基因。*abs*A 由 *abs*A1 和 *abs*A2 组成，编码双组分信号转导系统。这样的双组分信号转导系统与真核生物丝氨酸-苏氨酸蛋白磷酸系统相似。因此，在链霉菌次级代谢和分化发育中蛋白磷酸化信号传递起着重要作用。*abs*A 和 *abs*B 调控基因是通过调控途径专一基因而影响抗生素的合成。

3. 启动子

链霉菌启动子在序列结构上显示了复杂的多样性，至少存在 3 类链霉菌启动子。

① 与原核生物典型启动子－10 区和－35 区都类似的启动子；

② 仅与原核生物典型启动子－10 区相类似的启动子；

③ －10 区和－35 区与典型的原核生物启动子均不相似的启动子。

链霉菌之间的启动子序列和转录也很少相似性。链霉菌启动子的多样性还表现在－10 区和－35 区保守序列的间隔区长短差异较大在（7～24bp 之间）。有些链霉菌启动子序列与 *E. coli* σ^{70} 因子识别的序列相似，可以在大肠杆菌中表达，如 *S. coelicolor* 中 *act* Ⅱ-ORF1、ORF2、ORF3 与抗生素的外排有关，*act* Ⅱ-ORF1 和 ORF2/3 的－10 区和－35 区与大肠杆菌的相似。

4. 串联启动子

链霉菌基因含有串联或多个重叠启动子，这种结构和复杂的调控模式有关，在不同生长条件下负责基因分级表达，并受到时序的调控，如 *S. actuosus* 的 *nsh*A-p1 启动子、*S. aureofaciens* 质粒 pSA2201 中的 A1 启动子和 *S. coelicolor* 的 *sap*A 基因启动子直到各自机体生长后期才有活性。序列分析表明，链霉菌晚期启动子在转录起始位点附近的 DNA 序列上同 *sap* 启动子有一定的同源性。

链霉菌启动子转录起始位点并不都是单一的核苷酸，有时会把多重的邻接或邻近区域作

为起始位点。转录起始点通常离编码区上游 9～345bp，少数情况下转录起始点与翻译起始位点没有上游先导序列和明显的核糖体结合位点。通常，在起始密码子 ATG 或 GTG 第 5 到 12 个核苷酸上游，存在着与大肠杆菌 SD 序列类似的保守性序列，与 16S rRNA 的 3′端具有一定的互补性。

5. 蛋白分泌系统

(1) 蛋白分泌机制　大多数链霉菌的分泌蛋白的前蛋白中有 N 端信号序列，它们的分泌依靠 Sec 介导的分泌系统，从链霉菌中已克隆了 SecA、SecY、SecD、SecE 和 SecF 类似物。SecA 属于膜相关的转位 ATPase，它阻止分泌前蛋白形成三级结构前体，促进前蛋白定位于分泌的转位酶上。

作为能转位穿过生物膜的蛋白，必须以分泌的构型，至少有部分是由分子伴侣或热休克蛋白所介导。在 12 种链霉菌中发现了 2 个 *gro*EL 样基因（分别编码 HSP58 和 HSP56 蛋白）和 3 个 GRO EL 样蛋白，应答热休克 *gro* EL 样蛋白的 N 端与其他微生物相似。这暗示，链霉菌与大肠杆菌的分泌机制可能有类似之处。

链霉菌的翻译元件及分泌信息可以在大肠杆菌中起作用。将链霉菌 STI 和 STII 蛋白的基因连同核糖体结合位点和转录起始位点一起置于 pUC18 载体 *lac* 启动子下游，在大肠杆菌的细胞介质中可得到分泌表达。

(2) 信号肽　在链霉菌中分泌蛋白是以在 N 端含信号肽前体形式产生的，其 N 端在前体转位过程中或之后由信号肽酶分解下来。分泌信号肽的一级结构中没有同源性，但在二级结构中有相似的构型。链霉菌信号肽有以下共同的特点。

① 链霉菌信号肽比其他生物的信号肽长，链霉菌中有 56 个氨基酸长度的信号肽，平均为 29～31 个氨基酸，而大肠杆菌和真核生物一般只有 22～24 个氨基酸。链霉菌中多出的氨基酸位于亲水区（一般有 12 个氨基酸），比其他微生物的亲水区（4～8 个氨基酸）要长，有人认为这是链霉菌分泌机制在进化过程中的特异表现。还曾发现具有 70 个氨基酸的信号肽。

② 亲水区是链霉菌与其他所有生物信号肽的差别所在，该区含有 12 个氨基酸残基，带有很高的正电荷，富含精氨酸（30%），贫含赖氨酸（3%），可能是与链霉菌基因组高 GC 含量有关。该区在分泌过程中有重要作用。

③ 疏水区，以丙氨酸残基为主，参与信号肽插入脂质双层中。

④ 中性区，该区含有信号肽酶切割位点，一般为 Ala-X-Ala，含有一个脯氨酸残基，由于脯氨酸的结构不能使中性区形成 α 螺旋，有利于切割位点的暴露。

异源细菌信号肽可以被变铅青链霉菌识别或加工，如金黄色葡萄球菌噬菌体的葡萄糖激酶基因信号肽在变铅青链霉菌中表达人干扰素 α1，在上清液中可分泌表达 10^8U/L。利用链霉菌的信号肽序列可增加外源基因的表达。如革兰阴性细菌的 β-内酰胺酶可以在变铅青链霉菌靠自身信号肽分泌，然而如果用链霉菌信号肽取代，异源蛋白的分泌可增加 4～5 倍。

大多数分泌蛋白在 N 端有 20～30 个氨基酸残基是跨膜的分泌序列，尽管其一级结构很少保守，但均含有碱性 N 端（1 个或很多正电荷），中间 8～12 个疏水基团（非极性区）和非螺旋羧基端。实验证明 N 端碱性及中间非极性区对于蛋白跨膜转位是必需的，同时也会影响蛋白的合成。

（三）链霉菌基因表达系统

1. 链霉菌的载体

链霉菌的表达载体根据来源主要为质粒、噬菌体和柯斯载体。目前，根据在链霉菌中用

途的不同又可分为如下几种。

（1）高拷贝载体 pIJ101 是普遍采用的能在链霉菌中复制的高拷贝载体，经常用于鸟枪克隆，它们有广泛的宿主范围，通常加入硫链丝菌素（*tsr*）、红霉素（*ery*）、紫霉素（*vio*）和新霉素（*neo*）抗性基因作为选择性标记。pIJ702 是由 pIJ101 衍生的使用最广泛的质粒载体，具有两个选择性标记，*mel*（酪氨酸酶）和 *tsr* 基因，克隆在 *mel* 位点可使其失活，因不产生酪氨酸酶而导致不产生黑色素的表型，以此作为插入失活的标记。

（2）穿梭载体 由于在链霉菌中提取质粒的操作不像在大肠杆菌中那么容易，构建同时能在大肠杆菌和链霉菌中复制的穿梭载体是比较理想的，这样可以在大肠杆菌中完成重组质粒的构建，构建完成后转入链霉菌中进行表达，如 pHJL197、pHJL210(SCP2/pBR322)、pHJL302(SCP2/pUC19)、pZH1351 等。但是，有许多穿梭载体在链霉菌中不够稳定，尤其是来源于大肠杆菌的部分易丢失。外源基因也会表现出结构的不稳定。

（3）柯斯载体 含有柯斯位点的穿梭质粒载体，因包含外源 DNA 的容量比较，被广泛地应用于链霉菌基因文库的构建。

（4）噬菌体载体 链霉菌噬菌体载体大部分是由 ΦC31 衍生的，如 KC304、KC515/516 等。ΦC31 与细菌噬菌体在形态和普通遗传学方面均相似，能在大多数链霉菌中形成噬菌斑或形成溶源状态。它含有维持其溶源状态的阻遏基因及整合至宿主染色体的 *att*P 位点和一个黏性末端。

（5）分泌载体 链霉菌分泌多种蛋白酶，有许多蛋白酶基因已被克隆，这些基因可用于构建外源蛋白在链霉菌中的分泌表达载体。如由 *S. tendae* 产生的 β-淀粉酶抑制剂 tendamistat，用于构建分泌表达人 IL-2、人胰岛素原的载体。

（6）整合型载体 整合型载体可避免重组质粒的不稳定性，整合质粒可以在染色体上以单拷贝形式存在，不需要选择性压力而稳定存在。pSAM2 是一个 11 kb 元件，可以在许多链霉菌染色体上整合，其中 3.5kb 片段是整合重组所必需的，含有定点的重组酶。利用 pSAM2 的 3.5kb 片段构建了整合型载体 pPM927。另外，利用在变铅青链霉菌染色体中发现一段约 80kb AseⅠ片段对 ΦHAU3 显示抗性的现象，构建了在变铅青链霉菌中使外源基因能定点克隆的置换型载体 pHZ808，利用该载体可使克隆基因稳定地整合在染色体上，适合用于构建基因工程菌。

（7）高表达载体 为适应近年来基因组合生物合成研究新化合物的发展需要，利用天蓝色链霉菌中激活调节基因 *act*Ⅱ-ORF4 与 *act* Ⅰ基因启动子构建了整合型高表达载体 pCJR24、pCJR29，利用这些载体使重组改造后的红霉素生物合成基因表达水平提高了 10 倍。

2. 链霉菌基因转移的方法

基因引入链霉菌菌体的最常用方法是原生质体 DNA 的转化。然而，有一些链霉菌由于对外源 DNA 的限制修饰作用，运用此法时有一定困难。绕过链霉菌限制系统的方法有接合转移、电脉冲穿孔和噬菌体转染与转导等。

3. 宿主菌株

变铅青链霉菌是至今用于外源基因表达的唯一的链霉菌宿主菌，主要是因为大多数链霉菌有较强的限制修饰系统，而变铅青链霉菌缺乏明显的限制修饰系统，它能接受大多数启动子，还能分泌蛋白。如大肠杆菌的某些抗生素抗性基因、*tet*R、*lac*Z、*amp*C 等基因，金黄色葡萄球菌的 APH-Ⅲ磷酸转移酶基因、沙门菌的 β-内酰胺酶基因均在变铅青链霉菌中得到

表达。

现已有可利用的变铅青链霉菌重组缺失变株，能保证进入的质粒的稳定性。有研究证明，变铅青链霉菌能使蛋白糖基化，因而被认为是产生异源蛋白的有应用价值的宿主菌。

（四）影响链霉菌中基因表达的因素

1. 启动子对外源基因表达的影响

实验证明，启动子对外源基因的表达水平影响很大。如果使大肠杆菌的启动子序列接近其保守序列，则启动活性在大肠杆菌和链霉菌中均有所增加，但在大肠杆菌中增加得多，如果在启动子的－10 区和－35 区插入额外的 GC 碱基对，则启动活性在链霉菌中增加得更多。

在链霉菌中表达真核基因必须使用原核生物启动子。

2. 信号肽对外源蛋白分泌的影响

不同的信号肽 N 末端氨基酸序列正电荷数对基因表达的影响是不同的，有的正电荷数增加使外源蛋白分泌增加，但有的是减少而增加外源蛋白的分泌量。

有研究表明，信号肽和目的蛋白之间的距离对基因表达有影响。在信号肽和结构基因之间插入一段小肽，使两部分蛋白在空间上不产生位阻，可提高分泌和表达水平。

3. 密码子、SD 序列和终止子等对基因表达的影响

至今所测得的基因序列均证明，在链霉菌中翻译起始密码子为 ATG 或 GTG，终止密码子为 TAA 或 TAG。由于链霉菌含有高 GC 含量，约占 73%，尤其是第三位密码子 G'C 可达 90%以上，密码子利用倾向性对链霉菌基因表达是很重要的。为提高外源基因的表达量应注意其密码子使用情况，若符合链霉菌的偏爱性，则得到高效表达的可能性大。

实验证明，增加核糖体结合位点数有利于外源基因的表达。若将包括有 SD 序列的甘露聚糖酶信号肽（长度为 34 个和 43 个氨基酸）和纤维素酶 A 信号肽（长度为 27 个和 43 个氨基酸）插入木聚糖酶 A 结构基因上游（含有 RBS 位点）和启动子下游，木聚糖酶 A 的表达量提高 1.5～2.5 倍，RBS 之间需有 15～20 个核苷酸的间隔。

终止子在提高 RNA 聚合酶的利用率及增加 RNA 的稳定性中起重要作用，对外源基因在链霉菌中的表达也有影响，但是它的影响与所表达的蛋白有关。如将弗氏链霉菌 *aph* 的终止子插入到人 IFN-α2 基因下游，在变铅青链霉菌中表达，有 90%的转录在回文结构处终止，mRNA 转录效率提高，稳定性增加，使 IFN-α2 产量提高 4 倍。但是同一转录终止子插入到 tendamistat 基因下游，却对 tendamistat 产量没有影响。

4. 发酵条件对外源基因表达的影响

环境条件如温度、pH、溶氧对于宿主菌尤其是外源质粒的稳定性十分重要，同时也影响外源蛋白的有效表达。

发酵过程中保持最适 pH 是很重要的，过高 pH8.0～9.0 对基因表达产物不利。

高浓度蔗糖（34%）对产量有利，但研究表明，不是蔗糖本身的作用，而是蔗糖对水活性的影响。水活性在 0.98 和 0.96 之间，相当于蔗糖浓度为 270～280g/L。高浓度蔗糖（34%），或聚乙二醇浓度为 50g/L 和 100g/L 时，基因表达产量最高，可能是这种条件下酶活性比较稳定，或是水活性对酶的合成和分泌的渗透效应的影响。

温度对菌丝生长、质粒的稳定性和克隆基因产物的产生均十分重要。温度被认为是生化过程中最重要因素。

溶氧低对生长不利，但有利于基因表达，然而在氧受限的条件下质粒稳定性较差。溶氧在 10%～80%之间质粒均保持稳定，酶合成速率随着溶氧升高而提高。

（五）链霉菌表达系统的优缺点

链霉菌在作为外源基因表达体系有一定的优点，如：①链霉菌工业化培养条件成熟，适合于大规模产业化；②链霉菌基本为非致病菌，不产生内毒素。至今仅看到一例报道热吸水链霉菌是我国引起农民肺（过敏性肺泡炎）的主要病原之一；③链霉菌可以进行高密度培养，并能在相当一段时间保持其代谢活性，在稳定期细胞仍能合成蛋白次级代谢产物，因此可期望在稳定期仍能维持异源蛋白的产生；④链霉菌可分泌胞外酶，因此可将分泌信号肽序列纳入载体，使外源蛋白可以分泌胞外。在胞外分泌的异源蛋白能正确折叠的可能性比较大，而且较易被纯化；⑤链霉菌中有许多可利用的转录起始信号，利用来自链霉菌的能产生与分泌蛋白的分泌信号肽基因，可以较易地在链霉菌中进行基因产物的高表达。

在链霉菌中表达有生物活性的真核来源的蛋白至今有十余年的历史。研究实例显示，在链霉菌中表达的真核异源蛋白往往产量较低，远远达不到工业化生产水平。虽然链霉菌中的内源蛋白酶活性低于枯草杆菌，但链霉菌产生的蛋白酶对异源基因表达肯定是有障碍的，因此构建链霉菌蛋白缺失突变株对表达外源基因是很重要的。真核蛋白翻译后的加工在链霉菌中比较困难。现在已知生物体内许多活性多肽都是以酰胺化、糖基化形式存在的，链霉菌作为原核生物，其翻译后修饰能力有限，一般认为不能进行 C 端的酰胺化；链霉菌使外源蛋白糖基化的能力也有限。

链霉菌作为原核基因的表达及次级代谢产物的操作系统有其明显的优越性，而作为真核基因的表达系统目前看来还处在初级阶段，许多基础性研究工作有待进行。

第二节　外源基因在真核细胞中的表达

基因的体外重组和表达体系起始于大肠杆菌，但是基因工程技术的主要目的是生产真核细胞内有重要价值的蛋白质产品。因此，各种原核细胞表达系统在表达真核生物的基因时，或多或少具有这样或那样的缺点。所以，真核细胞表达系统的研究是今后基因工程研究的方向。

在真核细胞中常用的表达系统有酵母表达系统、昆虫表达系统、哺乳动物细胞表达系统和植物细胞表达系统等。下面将分别介绍酵母表达系统、昆虫表达系统，哺乳动物细胞表达系统在下一节单独介绍。

一、酵母表达系统

酵母是一种传统的食品和工业微生物，与人类生活关系密切。人类对酵母的认识和利用有数千年的历史，积累了大量有关酵母的知识，对酵母菌的一般生物学和分子生物学特性也已有很好的了解。酵母是单细胞的真核生物，有完整的亚细胞结构和控制严密的基因表达调控机制。它既能通过有丝分裂进行无性繁殖，也可以通过减数分裂实现有性繁殖，并具有生长繁殖迅速、营养要求简单和便于工业化大规模发酵培养的优点。

在 20 世纪 70 年代初建立了大肠杆菌表达系统之后，遗传学、生物化学、微生物学和分子生物学等学科的发展，为建立酵母表达系统创造了条件。1974 年，Clarck-Walker 和

Miklos 发现在大多数酿酒酵母中也存在质粒。这种质粒是双链 DNA，全长约 6300bp，每个二倍体细胞有 60～100 个拷贝，这就是后来被称为 2μ 环的质粒；随后在大肠杆菌中克隆得到了酵母的 *leu*2、*ura*3、*his*3 和 *trp*1 等基因。1978 年 Hinnen 将来自一株酿酒酵母的 *leu*2 基因导入另一株酿酒酵母，弥补了后者的 *Leu*2 缺陷，这标志着酵母表达系统的建立。自 1981 年 Hitzman 等用酵母基因表达系统表达了人干扰素以来，至今已有众多的外源基因在酵母系统中得到表达，其中包括获得美国 FDA 批准、用于人体的第一个基因工程疫苗——乙型肝炎疫苗等。鉴于酵母基因表达系统潜在的巨大经济价值，世界上许多国家都投入很大的人力物力对它进行研究开发。我国也在 1983 年首次用酵母菌表达了乙型肝炎病毒表面抗原基因。近 10 年来，人们已在除酿酒酵母以外的许多酵母菌中发展出了多个性能优良的表达系统，这些酵母表达系统各有特点，并都在表达外源基因的实际应用中取得了很好的效果，对推动生物工程的发展产生了很大的影响。酵母表达系统已在人类新基因功能的研究和新药研制等方面开辟了新的应用领域，表现了巨大的生命力。1996 年，在全世界科学家的通力合作下，完成了第一个真核生物——酿酒酵母全基因组的测序，这为人类对酿酒酵母的更深入研究及更广泛的利用打下了坚实的基础。

（一）酵母表达系统的组成

1. 酵母载体

（1）酵母载体的基本结构　尽管酵母中也有质粒（如酿酒酵母的 2μ 环质粒），但酵母载体一般都以大肠杆菌质粒为基本骨架，因为大肠杆菌转化方法简单、转化效率高，从大肠杆菌制备质粒 DNA 也比较方便，利用大肠杆菌系统构建酵母载体可以大大简化手续，缩短时间。因此，酵母表达系统的载体通常既能在酵母菌中进行复制，也能在大肠杆菌中进行复制，形成所谓酵母-大肠杆菌穿梭载体。

此外，酵母载体还具有以下一些构件。

① DNA 复制起始区　DNA 复制起始区是酵母细胞核内 DNA 复制起始复合物的结合位点，赋予酵母载体在每个细胞分裂周期的 S 期自主复制的能力。这种序列通常来自酵母菌的天然 2μ 环质粒的复制起始区及酵母基因组的自主复制序列（autonomously replicating sequence，ARS）。

② 筛选标记　筛选标记是载体转化酵母后筛选转化子时必需的构件。酵母表达系统中常用的筛选标记有两类：一类是酵母宿主为营养缺陷型，如 Leu2、Ura3、His3、Trp1 和 Lys2 等，其筛选标记就是营养合成代谢途径中相应的 *leu*2、*ura*3、*his*3、*trp*1 和 *lys*2 基因；另一类是显性筛选标记，可用于转化野生型酵母菌，如 G418 和放线菌酮（cycloheximide）等。

③ 整合介导区　这是与宿主基因组有同源性的一段 DNA 序列，它能有效地介导载体与宿主染色体之间发生同源重组，使载体整合到宿主染色体上。这种同源重组的过程主要有两种形式：单交换整合与双交换整合。单交换整合（single cross-over integration）的结果通常是在整合转化子染色体的整合位点附近又增加了一份同源序列的拷贝，所以已整合上去的载体有可能因这两份同源序列之间的重组又从染色体上切割下来。但自然发生的同源重组频率非常低，所以单交换整合转化子一般还是相当稳定的。双交换整合又称替换或置换（replacement 或 transplacement），是整合载体的一部分通过在两个不同位点与染色体发生同源重组而整合入酵母基因组，并同时置换下这两个位点间一段染色体 DNA。双交换整合的结果不会在整合位点附近形成同源序列的重复，避免了再次发生同源重组的可能性。所以，双交换整合转化子是非常稳定的。可通过选择特定的整合介导序列，人为地控制载体在宿主染

色体上的整合位置与拷贝数。一般地说，酵母染色体的任何片段都可作为整合介导区，但最方便、最常用的单拷贝整合介导区是营养缺陷型筛选标记基因序列。酵母基因组内的高拷贝重复序列（如 rDNA、*Ty* 序列等）则可作为多拷贝整合介导区。

④ 有丝分裂稳定区　在细胞有丝分裂时，游离于染色体外的载体能否有效地、平均地分配到子细胞中是决定转化子稳定性的重要因素之一。有丝分裂稳定区能帮助载体在母细胞和子细胞之间平均分配。常用的有丝分裂稳定区来自于酵母染色体的着丝粒片段。此外，来自酵母 2μ 质粒的 STB（stability）片段也有助于提高游离载体的有丝分裂稳定性。

⑤ 表达盒　表达盒（expression cassette）是最重要的构件。它主要由转录启动子和终止子组成。如果需要外源基因的表达产物分泌，在表达盒的启动子下游还应该包括分泌信号序列。由于酵母对异种生物的转录调控元件的识别和利用效率很低，所以表达盒中的转录启动子、终止子及分泌信号序列都应该来自酵母本身。

启动子是表达盒中的核心构件，长度一般在 1～2kb 之间。启动子下游有转录起始位点和 TATA 序列，普遍性转录因子能识别转录起始位点及 TATA 序列，形成转录起始复合物。转录起始复合物决定了一个基因的基础表达水平。启动子上游还有各种调控序列，包括上游激活序列（upstream activating sequence，UAS）、上游阻遏序列（upstream repression sequence，URS）和组成型启动子序列等。一些调控蛋白可与之相结合，并和转录起始复合物相互作用，以激活、阻遏等方式影响基因的转录效率。

终止子是决定 mRNA 3′-末端形成效率的重要元件。与高等真核生物一样，酵母中 mRNA 3′-末端的形成，也经过前体 mRNA 加工和多聚腺苷酸化反应。不同的是，在酵母中这些反应发生在基因 3′-端的近距离内，并紧密偶联。所以，酵母基因的终止子一般不超过 500bp。

分泌信号序列也称信号序列（signal sequence），编码前体蛋白上 N 端一段 17～30 个氨基酸残基的分泌信号肽。分泌信号肽的作用是引导分泌蛋白在细胞内沿着正确的途径转移到胞外，这对于分泌蛋白的翻译后加工和生物活性都有重要意义。一般来说，酵母细胞识别外源分泌蛋白的信号肽进行蛋白质的输送和分泌表达产物的效率较低，所以需要依赖酵母本身的分泌信号肽来指导外源基因表达产物的分泌。常用的酵母分泌信号序列有 α 因子的前导肽序列、蔗糖酶和酸性磷酸酯酶的信号肽序列。其中 α 因子的前导肽序列指导表达产物最为有效，在各种酵母菌中的适用范围最广。

（2）酵母载体的种类　酵母表达载体可以根据载体在酵母中复制形式、载体的用途、载体表达外源基因的方式等来分类。鉴于载体在酵母细胞中的复制形式是酵母载体最重要的特性，以此为标准将酵母载体进行分类，一般可以把它们分为 YIp、YRp、YCp、YEp 和 YAC 5 类。

① YIp（yeast integration plasmid）型载体　这是一种整合型载体。它不含酵母的 DNA 复制起始区，因此不能在酵母中自主复制；但它带有整合介导区，可通过同源重组而整合入酵母基因组并随同酵母染色体一起复制。在整合介导区用限制性内切酶将 YIp 型载体切成线形分子，可以大大增加整合效率。酵母整合型载体因其转化子的高度稳定性而被广泛应用。但它有 2 个不足之处，一是转化频率很低（$<10^2$ 个转化子/μg DNA），二是它的整合拷贝数一般都很少（1～2 个拷贝/细胞），因而其转化子对外源基因的表达量相对较低。

如果用在酵母染色体上以多拷贝形式存在的 DNA 片段如 rDNA、Ty 序列等作为整合介

导序列构建整合载体，就可以大大提高整合的拷贝数和外源基因的表达水平。这类整合载体被称为 YMIp（yeast multiple-integration plasmid）型载体。

② YRp（yeast replicating plasmid）型载体　这是一种自主复制型载体，它含有酵母基因组的 DNA 复制起始区，因此能在酵母染体外自主复制。YRp 质粒的特点是转化效率高，1μg DNA 分子可形成 $10^{3\sim4}$ 个转化子，并且每个细胞的质粒拷贝数可高达上百个。但是这种类型的载体在细胞分裂时很难在母细胞与子细胞之间平均分配，而且大多滞留在母细胞内，因此即使在有选择压力的条件下，随着转化细胞不断地分裂繁殖，子代细胞中的 YRp 质粒拷贝数也会迅速减少，最终导致整个群体的平均拷贝数变得很低（只有 1～10 个拷贝/细胞）。如果在没有选择压力的条件下培养，丢失了载体的细胞更会以每世代高达 20%的速率累积。因此，YRp 质粒是一种较好的建库载体，也可作为实验研究的表达载体，但难以用于工业生产。

③ YCp（yeast centromeric plasmid）型载体　这种载体是在 YRp 质粒的结构基础上，增加了一段来自酵母染色体着丝粒的 DNA 片段而形成的。酵母着丝粒的存在可以使这种载体在细胞分裂时能像宿主染色体那样在母细胞与子细胞之间平均分配，所以，其转化子细胞每世代丢失质粒的频率不到 1%，表现出高度的质粒稳定性。然而，与此同时其 DNA 复制受到严格控制，每个细胞中的质粒拷贝数只有 1～2 个。YCp 质粒常用于构建基因文库，特别适用于克隆和表达那些多拷贝时会抑制细胞生长的基因。

④ YEp（yeast episomal plasmid）型载体　这是一种酵母附加体型载体，含有酿酒酵母 2μ 质粒与 DNA 复制有关的部分或全部序列，因此有很高的转化效率（约 $10^{3\sim4}$ 转化子/μg DNA）。野生型的 2μ 质粒在酵母细胞中非常稳定，并且每个细胞中的拷贝数可以高达 60～100。这是因为野生型 2μ 质粒存在 FLP-FRT 位点特异重组系统，使它可以具有超越染色体 DNA 复制周期而增加 DNA 复制的机会。此外，2μ 质粒的复制除了需要 *ori-stb* 区外，还需要自己编码的 *rep*1 和 *rep*2 基因的配合。最近研究表明，在 2μ 质粒的 *Sna* BI 位点附近为一非必要区。将构建酵母载体的所有其他构件都插入这个位点，就能保持 2μ 质粒的完整性，从而使其成为一个高稳定、高拷贝的 YEp 型载体。

⑤ YAC（yeast artifical chromosome）型载体　即酵母人工染色体型载体，它具有酵母染色体的主要构件，包括酵母染色体自主复制序列（ARS）、着丝粒序列（CEN）和端粒序列（TEL）。导入酵母细胞的 YAC 载体以线性双链 DNA 形式存在，具有高度的遗传稳定性。但其复制受细胞分裂周期的严格控制，一般每个细胞中都只有单拷贝。YAC 载体上可以插入大到 500kb 的 DNA 片段，可用于高等真核生物基因的克隆。近年来的研究表明，YAC 载体或经过修饰的 YAC 载体还可以导入哺乳动物细胞或人的细胞。

2. 宿主

酵母种类繁多，已知有 80 个属约 600 多种、数千个分离株，是一类巨大的、很有应用前景的生物资源。总的来说，作为基因表达系统宿主的酵母应该具有以下条件。

① 安全无毒，不致病。极大多数的酵母菌是安全的，只有极少数是致病菌或条件致病菌。

② 有较清楚的遗传背景，容易进行遗传操作。

③ 容易进行载体 DNA 的导入。由于 DNA 转化技术的不断发展，只要对转化条件进行优化，多数酵母菌可以取得较高的转化效率。

④ 培养条件简单，容易进行高密度发酵。

⑤ 有良好的蛋白质分泌能力。

⑥ 有类似高等真核生物的蛋白质翻译后的修饰功能。

在各类酵母中，酿酒酵母（*Saccharomyces cerevisiae*）最符合上述条件，因此最早成为基因表达系统的宿主，现已被广泛用来表达各种外源基因。用酿酒酵母做宿主表达的乙型肝炎疫苗、人胰岛素和人粒细胞集落刺激因子等基因工程产品均已正式上市。

酿酒酵母也有不足之处：①发酵时会产生乙醇，乙醇的积累会影响酵母本身的生长，因此较难进行高密度发酵；②分泌蛋白质的能力较差；③对蛋白质的糖基化修饰不够理想，与高等真核生物相比所形成的糖基侧链太长，这种过度糖基化可能会引起副反应。面对酿酒酵母上述问题，人们不仅对其进行遗传改造，改善其特性，还试图从酵母菌这一大家族中寻找更好的宿主。

近年来，人们已在除酿酒酵母以外的许多酵母菌中发展出了多个性能优良的表达系统，这些酵母表达系统各有特点，并都在表达外源基因的实际应用中取得了很好的效果。巴斯德毕赤酵母（*Pichia pastoris*）表达系统就是其中成功的一个，已被广泛应用。该酵母表达系统可以使发酵密度达到很高的水平，分泌蛋白质的能力强，糖基化修饰功能更接近高等真核生物，弥补了酿酒酵母的不足；但是也有其缺点：①分子生物学的研究基础差，要对其进行遗传改造较困难；②不是食品微生物，发酵时要添加甲醇，所以要用它来生产药品或食品还无法被广泛接受；③发酵周期较长。可见构建新的酵母特异表达系统是很有吸引力的工作，但是要建成一个系统并被广泛接受需要做大量的工作。

3. 载体转化

酵母有结构复杂的细胞壁，最早人们应用原生质体法进行载体 DNA 的酵母转化。但酵母细胞原生质体化的程度难以控制，而控制不好就会大大影响转化效率，因此此法不容易掌握。另外，原生质体转化也昂贵、费时。

有人发现先用一价碱性阳离子 Cs^+、Li^+ 溶液处理酵母细胞，然后进行 DNA 转化，能明显地增加外源 DNA 的转入，首次实现了完整酵母细胞的 DNA 转化。这种方法的转化效率较原生质体法低，但也能达到 10^3 个转化子/μg DNA，足已满足一般的应用要求。此外，这个方法简便、容易掌握，所以很快被广泛采纳。在此基础上，又建立了酵母转化一步法，使酵母转化的方法变得越来越简单，特别适用于转化处于静止期的酵母细胞。

最早用于植物细胞的 DNA 转化的电穿孔法和粒子轰击法，后来证明也能用于酵母细胞的转化，而且成功率较高，常常被用来进行一些新的酵母宿主细胞的 DNA 转化。但是这种 DNA 转化技术需要特殊的物理设备，推广有一定难度。

（二）酵母表达系统的应用策略

酵母表达系统主要用来表达外源基因，获得可用作药品和食品添加剂的基因工程产品。因此，外源基因的表达水平和产物质量是研究和利用这一表达系统的首要问题。

1. 提高外源基因在酵母中的表达效率

外源基因能否高效表达直接关系到这些表达产物是否具有开发价值。外源基因产物应用时的剂量及原有生产成本不同，则所需的表达量不同。如细胞因子在临床上的用量一般只有每剂几微克到几十微克，而此前无法获得这种产品，因此每升发酵液能表达几十毫克已经有意义。人血清白蛋白临床上的用量在每剂 5～10g 的水平，传统生产成本又较低，所以工程菌的表达水平必须达到 10g/L 以上。

人们已经在基因表达的多级水平上进行改进，有效地提高了外源基因的表达水平。

(1) 转录水平控制 外源基因在酵母中的表达和基因的转录水平有密切的关系，筛选高效的启动子就显得十分重要，强启动子还可以表达多种基因。和糖酵解有关的酶在酵母细胞中的含量特别高，所以筛选和克隆有关的基因首先成为大家的目标。磷酸甘油酸激酶、甘油醛磷酸脱氢酶分别可占细胞可溶性蛋白质的5%左右，其启动子PGK1和启动子GAPDH（或GAPl）是目前所知最强的启动子。利用这两个启动子已高效表达了许多有用的外源基因，如人干扰素、牛凝乳酶、免疫球蛋白和HIV抗原、乙型肝炎病毒表面抗原及核心抗原、抗胰蛋白酶、人SOD和人表皮生长因子等。但用组成型启动子表达外源基因也存在问题，其基础表达水平很高，从工程菌在培养基中繁殖一开始外源基因就有很高的表达，表达的产物可能会严重影响酵母细胞的生长，并且还会造成低表达突变体生长优势的选择压，以致最终并不能得到高表达。

酵母菌也有很多表达过程受严格控制的基因，这些基因的调控型启动子可用于指导外源基因表达。让工程菌首先在这些启动子关闭的条件下生长，当细胞生长达到一定密度，再改为使这些启动子打开的生长条件，高密度的工程菌细胞可以使外源基因大量表达，此时表达产物对酵母细胞的毒害作用已经不明显了。但是，调控型启动子即使在启动子打开的条件下，转录效率也没有PGK1和GAPDH启动子强，因而限制了外源基因的表达水平。

近一步研究表明，染色体GAPDH基因启动子和表达质粒上长启动子可能有相同的增强子序列，它们的工作需要相同的转录因子，而细胞内的转录因子数量有限，转录因子的竞争既影响了外源基因的表达，也影响了内源GAPDH基因的表达。因此，将启动子上游激活序列（UAS）和GAPDH的短启动子融合，得到杂合启动子。这类杂合启动子同时具有可调控及强转录启动功能的优点，并能高效表达外源基因。

载体在染色体上的整合位点对外源基因的表达也有很大的影响。和高等真核生物一样，酵母染色体上也存在基因转录的沉默区，主要在端粒附近及HML和HMR区。不过酵母染色体上对基因表达有沉默效应的区域不大，因此不必过多地考虑。

(2) 表达载体的拷贝数和稳定性 表达载体在细胞中的拷贝数对外源基因在酵母中的表达有明显的影响。酿酒酵母和其他一些酵母有多拷贝的内源质粒，以这类质粒为基础可以构建高拷贝质粒表达载体。由于游离质粒构建的载体在细胞中的稳定性尚未完全解决，以及具有内源质粒的酵母菌株并不多，因此常用整合性载体，但是所有酵母菌株都有rDNA，多用rDNA这样的多拷贝基因来介导载体的整合。有很多种酵母菌利用rDNA片段介导了表达载体的多拷贝整合，并使外源基因得到相当高水平的表达。但有两个因素会造成整合拷贝不稳定：一是构建的整合载体的大小；二是rDNA片段的序列。

实验证据表明以rDNA介导的整合载体的大小应该控制在9～10kb，这相当于染色体上一个rDNA单位的大小。而在整合载体中被选用的rDNA片段的序列也会影响其在细胞中的稳定性。在酵母rDNA单元中有两个HOT区和一个TOP区，前者能刺激rDNA间的重组，而后者则能抑制这种重组。如果所取的rDNA片段只有HOT区，就容易发生重组，整合拷贝容易丢失。如果所取的rDNA片段中既有HOT区，又有TOP区，或者两者都不存在，整合拷贝间不容易重组，整合拷贝就很稳定。在适当的选择条件下，以染色体上单拷贝DNA片段为介导，也可以构建成稳定的多拷贝整合载体。

(3) 其他因素 还有一些因素应该考虑，其中包括：采取提高翻译效率的措施，如优化翻译起始区前后mRNA的二级结构、在外源基因中尽量选用酵母偏爱的密码子等；避免表达产物在细胞内的降解；选择或改造宿主，如采用二倍体宿主、采用酿酒酵母以外

的酵母菌作为宿主；优化工程菌发酵工艺，如提高发酵密度、控制发酵阶段和发酵时间等。

2. 提高外源基因表达产物的质量

外源基因在酵母表达系统中高表达，使本来很难得到的生物活性蛋白质有可能开发成为新药。为了保证这类新药的疗效，避免可能产生的毒副反应，还要考虑表达产物的可靠性，即酵母表达的产物在结构上尽可能与天然产物一致。随着蛋白质结构检测手段的发展，过去检测不到的结构变化现在能检测到了，有时发现这种变化还十分可观。因此表达产物的可靠性日益引起人们的重视。影响表达产物可靠性的因素很多，其中包括：①外源基因在表达系统中的遗传稳定性；②不同生物在蛋白质合成、加工和修饰功能上的差别；③产物在短时间的过量表达等。

在表达系统中突变事件经常会发生。如果表达载体以多拷贝形式存在，大多数突变并不会对表达产物产生影响。但有时这种突变不仅影响产物的结构，同时还会使带突变体的酵母细胞具有生长优势，在高密度发酵条件下，随着发酵时间的延长，这种突变表达产物的比例就会逐渐增加。

表达产物要成为有功能的蛋白质还需经过多种加工和修饰过程，包括去除 N 端甲硫氨酸残基、氨基端乙酰化、羧基端修饰、糖基化及多亚基的聚合等。

实验表明，酵母具有使胞内表达的外源基因产物在氨基端进行乙酰化的能力。酵母细胞内有甲硫氨酰氨肽酶，它可以使胞内表达的蛋白质去除氨基端的甲硫氨酸残基，但去除往往不完全。此时，可以采用泛素融合方式来表达外源基因，不仅提高了表达水平，还可利用泛素水解酶的作用完全避免了表达产物氨基端甲硫氨酸残留的问题。

对蛋白质羧基端的甲基化、豆蔻酰化和法呢酰化 3 种修饰作用是膜蛋白着膜表达所必需的。在酵母中，已成功地表达了多种有功能的动、植物膜蛋白，说明酵母细胞内也存在和动、植物一样的羧端修饰酶。另外，酵母菌有较强的亚基装配能力，乙型肝炎病毒表面抗原亚基间、哺乳动物钠钾 ATP 酶的 α 亚基、β 亚基间及人胚胎血红蛋白的 α 亚基、β 亚基间都能在酵母细胞中成功地表达和装配。大多数用作药物的蛋白质，其天然状态一般都是糖基化的，糖基化与蛋白质的生物活性有着密切的关系。虽然有一些天然产物是糖蛋白而基因工程得到的未经糖基化的产物，也有很好的生物活性，但是未经糖基化的产物可能产生新的抗原性，这也是值得注意的问题。酵母细胞的 *N*-糖基化与真核生物一样，但其合成的糖链多为甘露糖残基，并会产生过度糖基化现象；而酿酒酵母的 *O*-糖基化也和高等真核生物的不同。这类糖基化上的差别可能也会影响生物活性或引起过敏反应。

酵母还可以对很多外源基因进行分泌表达。分泌表达可以提高表达水平，简化表达产物的纯化步骤，还可以使表达产物通过分泌途径完成一系列的加工和修饰，这对保证表达产物的天然活性是十分重要的。酵母表达的单链抗体、人胃脂肪酶、成纤维细胞胶原酶、IL-8 都具有与天然产物相同的活性。组织蛋白酶 B 不能在大肠杆菌和昆虫细胞系统中表达，但是可以在酵母细胞中分泌表达出活性很好的产物，说明酵母细胞有一定能力通过分泌途径进行蛋白质的翻译后修饰。有时，一个外源基因不能在酵母中分泌表达天然性好的产物，但是，适当改变载体的构建条件，情况可以得到改善。有实验表明分泌表达水平不仅和启动子的强度有关，而且还和所用载体的类型有关。

此外，产物合成的速度应该和折叠修饰速度相匹配，细胞中蛋白质的折叠修饰能力也是可以提高的。

二、昆虫表达系统

昆虫因可大量繁殖、能完成表达产物的一系列翻译后加工修饰、载体和表达产物对人危害小等特点，正日益成为一个非常有用的生物技术表达宿主，成为新兴的生物制药资源，对于人类疾病治疗和疫苗的开发，改善人类的健康和提高生活质量提供了良好的前景。

20 世纪 80 年代初首次建立了苜蓿尺蠖核型多角体病毒（AcNPV）-秋黏虫细胞杆状病毒表达系统（baculovirus expression vector system，BEVS），1984 年又建立了家蚕核型多角体病毒（BmNPV）-家蚕表达系统。以下将分别介绍。

（一）杆状病毒表达系统

杆状病毒科是病毒粒子呈杆状、内含单分子超螺旋闭环双链 DNA 基因组（88～166kb）的一类病毒，它主要对昆虫显示病性。杆状病毒的基因组很大（约 130kb）、限制性酶切位点太多，难于直接对其进行常规的克隆工作。一般采用将外源基因克隆到专门设计的转移载体，与病毒 DNA 共转染进细胞（细菌）经同源重组途径，再空斑筛选、纯化而得到重组病毒。杆状病毒科的两个代表 AcNPV 和 BmNPV 的全部序列已经测定，其大小分别为 134kb 和 128kb。两者的主要基因排列顺序基本一致，核酸同源性在 90%以上。

1. 杆状病毒表达系统的特点

（1）强启动子　选用病毒基因组强启动子，使外源基因高效表达。常利用病毒晚期基因启动子，外源基因表达在大部分病毒粒子形成后才启动，此时细胞的转录、翻译机构已被病毒所控制，细胞对表达产物毒性的敏感性明显降低，仍可高水平表达。但表达时相很迟（感染后 72h），不利于产物的翻译后加工。也可利用病毒极早期基因启动子，可将外源产物表达峰值提前，为即时和受控表达外源基因产物提供了可能；可实现持续表达、较高表达水平。非常适用于某些后加工要求高的基因和杆状病毒杀虫剂，已被广泛应用于外源基因的提早表达和建立稳定转化的细胞株。

（2）可容纳较大、较多的外源基因　杆状病毒的核衣壳和基因组较大，一次可以容纳大至 12kb 的外源基因片段。借助于多元表达载体或几个不同的重组病毒共同感染可以同时表达多个外源基因。

（3）可进行翻译后加工修饰　作为较高等真核生物，昆虫细胞能完成一系列翻译后加工修饰工过程，绝大部分外源基因产物均显示良好的生物活性。

（4）安全性　BEVS 被认为遗传学上是安全的表达载体。杆状病毒对脊椎动物无病原性，也不能在哺乳动物细胞内进行病毒 DNA 复制、转录、长期滞留或整合到哺乳动物细胞基因组内。外源基因通过同源重组将多角体蛋白基因敲除掉后，重组病毒不能表达包含体蛋白。这不仅提供了选择标记，而且重组病毒不能像野生型病毒那样在环境中长期留存，对宿主昆虫最主要的感染途径——经口感染也不具有病原性。

2. 病毒载体

杆状病毒在进化过程中保留或获得了一些基因以利于其繁殖和传播，如胱氨酸蛋白酶和几丁质酶，前者在病毒感染晚期大量表达，分解虫体使其容易腐败，利于病毒的传播，后者的大量表达可以破坏表皮几丁质结构，使虫体液化，使病毒大量扩散以利于再次感染。这些基因能够引起表达产物的降解，在活体表达中更能引起虫体组织的溃败而提前死亡，降低了表达量，增加了纯化难度。删除这些基因如胱氨酸蛋白酶、几丁质酶，可提高表达产物的稳定性；删除蜕皮甾体 UDP-葡萄糖基转移酶（Egt）可提高杆状病毒的杀虫效果，缩短宿主

的存活时间。有报道用半乳糖苷酶替代 BmNPV 中的胱氨酸基因，发病蚕体不易腐败，外源基因表达量得到了 40%的提高。

蜕皮甾体葡萄糖基转移酶（*egt*）基因的编码产物能使蜕皮激素失活，推迟幼虫入眠，引起幼虫持续取食和生长，增加了感染幼虫的病毒繁殖量，利于病毒的进化。将 BmNPV 中的 *egt* 基因敲除后，幼虫体重减轻，死亡时间提前。故 *egt* 被敲除的修饰病毒可用于生产生物杀虫剂。

3. 杆状病毒表达系统（BEVS）

BEVS 系统有重组率低、筛选困难、工作周期长等不足之外。针对上述缺点，围绕着扩大应用范围、缩短工作周期已进行了如下改进。

（1）优化筛选方法　利用水母绿色荧光蛋白（green flurescent protein，GFP）及其突变体作为报告基因，具有检测手段简单、表达产物稳定和发光反应不需要其他蛋白和辅助因子及底物的参与等优点，已将其引入 BEVS 作为重组病毒的筛选标记。利用单纯疱疹病毒Ⅰ型的胸苷激酶基因（HSV1-TK）能催化核苷酸类似物成为 DNA 复制的抑制剂这一特性，开发了一种新的重组病毒筛选方法。核苷酸类似物（Ganciclovir）直到 100μmol/L 时并不影响 sf-9 细胞活力和 AcMNPV 野生型病毒的复制，而将病毒的极早期启动子（IE-1）启动的 HSVl-TK 基因通过常规的同源重组技术引入 AcNPV 基因组获得重组病毒 AcNPVIE-1-TK，其复制受到 Ganciclovir 剂量依赖性抑制：在 2μmol/L 时即可明显抑制，在 100μmol/L 时复制完全受阻。因此，将连接有外源基因的转移载体和 AcMNPVIE-1-TK 病毒 DNA 共转染时，将子代病毒接种在含 50μmol/L Gandclovir 的细胞培养基生长的细胞进行连续培养，只有 TK 基因已经被外源基因所替换的病毒才能良好地生长。因此，只需一轮纯化即可获得重组病毒，简单快捷，成功率高达 85%以上，且获得的重组病毒完全与常规相同。

（2）提高重组率　在多角体基因的下游存在一个病毒复制的必需基因 ORF9（1629bp），编码核衣壳的一个组分；而在其上游有一个非必需的基因 ORF603。通过体外突变得到一个在 ORF603 和 ORF9 各有一个 *Bsu*36 Ⅰ酶切位点的突变型病毒株 BacPAK6。将用 *Bsu*36 Ⅰ酶切的 BacPAK6 DNA 与带有多角体启动子、外源基因和复制必需基因 ORF9 的重组转移载体共转染昆虫细胞，只有外源基因连同复制必需的基因重组到病毒 DNA 后才能存活，重组率可高达 80%或更高。此外，通过转移载体将 *Sca* Ⅰ酶切位点引入 AcNPV 的病毒基因组得到重组病毒，该重组病毒重组率达到 95%以上，每微克病毒 DNA 可得到 2×10^5 个重组病毒。这种方法非常适合于建立真核生物基因的 cDNA 表达文库和多肽文库。

（3）建立稳定表达系统　在绝大多数情形，用昆虫细胞生产野生型杆状病毒或用 BEVS 生产外源蛋白质，所有细胞都因病毒感染而死亡。因为外源基因的表达是间歇的，常在感染的晚期，高效表达的重组产物的后加工不完全。另外经多次连续传代后，病毒对昆虫细胞的感染力降低，外源基因丢失，产生所谓的缺损病毒，因此连续培养的有效表达时间仅一个月，影响了表达效率。Jarvis 应用病毒的极早期启动子构建了一个稳定表达载体，可连续表达 100 代（1 年）以上，β-gal 的转化细胞的表达量为 2mg/L 细胞，仅及用 BEVS 表达的 1%。与传统的 BEVS 相比较来看，BEVS 适于表达那些非分泌型的、糖基化要求不高的表达产物；而稳定表达系统适于表达分泌型的、加工要求复杂的外源蛋白，将后者与无血清培养基、无蛋白培养基的开发相结合，将是传统 BEVS 的一个有力补充。糖基化是翻译后加工过程的重要环节，昆虫细胞的糖基化途径有别于高等的哺乳动物细胞。1996 年 Jarvis 等将糖基化修饰基因置于 IE-1 启动子控制下，建立稳定化表达载体，修饰了昆虫细胞的 *N*-糖

基化途径。

(4) 构建穿梭载体　包括酵母-昆虫细胞穿梭载体、大肠杆菌 *E. coli*-昆虫细胞的穿梭载体。前者主要组分有：酵母的自主复制序列（ARS）；CEN，起有丝分裂中着丝粒的作用，保证质粒 DNA 能稳定地分配到子代酵母中去；反选择标记 SUP4-0；URA3 为选择标记基因。这种重组病毒既可在酵母中复制，也能在昆虫细胞中复制。用外源基因替代 SUP4-0 基因，此重组穿梭载体能在含刀豆氨酸（canavanine）培养基的酵母中生长，抗 canavanine 的酵母克隆即为含外源基因表达重组病毒，最快可在 10～12d 之间获得。

E. coli-昆虫细胞穿梭载体（bacmid）是迄今为止最为方便、快捷的重组杆状病毒筛选系统。bacmid 包含有 mini-F 复制子、卡那霉素抗性筛选标记和细菌转座子 *Tn7* 的靶序列 *att*Tn7 及 *lacZ* 片段，能像质粒一样在 *E. coli* 中复制，也能像病毒一样感染昆虫细胞。采用供体质粒取代转移载体，供体质粒用于提供外源基因，其设计为含抗庆大霉素基因和多角体蛋白基因启动子启动的外源基因，两侧分别为 *Tn7* 的左臂和右臂。同时为了产生位点特异的转座，还需一个辅助质粒（该质粒为四环素抗性筛选标志）提供转座蛋白。将 bacmid、供体质粒、辅助质粒转化同一个菌株，由辅助质粒产生转座蛋白，将抗庆大霉素基因和多角体蛋白启动子控制的外源基因转座到 bacmid 的 *attTn7* 序列上，同时破坏了 *lacZ* 的阅读框架，这样只要将转化上述质粒的菌株在含卡那霉素、四环素、庆大霉素及 X-gal 和 IPTG 的 LB 培养基上筛选白斑即可得到复合 bacmid，将复合 bacmid DNA 纯化后转染昆虫细胞，就得到了纯的重组病毒。该系统是目前 BEVS 中最为方便、快捷的系统，整个重组病毒筛选的周期可缩短为 7～10d，并可以连续筛选，非常适合构建 cDNA 库及用于蛋白质工程。同时，供体质粒较常规转移载体小，可以构筑较多的限制性酶切位点以利于外源基因的插入；且能适合在其他真核生物表达系统中进行转座以表达外源基因。

(5) 体外重组表达系统的构建　利用噬菌体 P1 特异性的重组系统 Cre-lox 可构建体外重组表达载体系统。主要原理如下：噬菌体 P1 含有一个 Cre 重组酶，能特异性识别 loxP 序列，并与之发生重组。将人工合成的 loxP 序列通过同源重组转移进 NPV 基因组，产生重组病毒 NPV（$loxP^+$）；当含 LoxP 的转移载体和 NPV（$loxP^+$）在体外经 Cre 重组酶的催化可发生位点特异的重组，转染细胞后很容易筛选到重组病毒。这个系统仅需一个步骤就能将外源基因从病毒基因组中分离出来，适用于 cDNA 文库的构建。

4. 昆虫细胞的培养

昆虫细胞体积大，形状较不规整，容易结团，加之病毒感染后细胞膨大 1.5 倍左右，这些因素增加了昆虫细胞大规模培养的难度。目前，已经建立的大规模昆虫细胞培养形式有贴壁培养、悬浮培养和固定化培养等。其中，贴壁培养又可分为滚瓶培养和微载体培养两种；悬浮培养又可分为摇瓶和转瓶，气升式反应器和搅拌式生物反应器。固定化培养是近来发展的一项昆虫细胞大规模培养新技术，具有两大优点：可有效降低剪切力对细胞的伤害；细胞与培液可有效分离，简化了操作步骤，降低了染菌概率。它又可分为微囊化培养和堆积床法培养，其中堆积床法培养综合了各种培养法的优点，可提供高培养表面，同时又能保证充足的供氧，获得最高的细胞培养密度。

5. 杆状病毒表达系统的应用

(1) 在蛋白质结构和功能上的研究　BEVS 表达的外源蛋白能进行同源复合物和异源复合物的超分子装配。如血小板生长因子可高效地进行二聚化，只有很少一部分保持单体形状，而 HBsAg 可以装配成更大的同源寡聚体颗粒；而用 BEVS 表达的全脊椎灰质病毒编码

的蛋白分解酶能加工多聚蛋白成 vp0、vp1、vp3，这些产物装配成异源寡聚体，成为稳定的无感染性的脊椎灰质类病毒类似颗粒。BEVS 表达产物的糖基化较简单，比在哺乳动物细胞所表达的异质性减少许多，这对高分辨率技术如 X-射线晶体结构分析和核磁共振来说是求之不得的优点，加之 BEVS 表达效率高，易得到晶体，许多实验室将 BEVS 用于蛋白质晶体学的研究。

（2）真核多肽文库（surface display）技术　该技术可以从一个已构建的库中筛选出结合蛋白，是研究配体-受体、抗原-抗体及多肽药物分析的一个利器。杆状病毒的膜蛋白 gp64 裸露在病毒粒子的囊膜表面，而且即使插入外源的多肽片段也不会使该蛋白失去功能，是一个理想的外源蛋白定位位点，已用于构建杆状病毒的多肽文库。另一个可行的多肽定位位点是用杆状病毒表达的膜结合蛋白。HIV-1-gp41 的抗原决定簇“ELDKWA”能被人中和性抗体 mAb2F5 所识别。流感病毒的血凝素表达在重组病毒感染后的昆虫细胞表面。为了增强该抗原决定簇和抗体的特异性结合力，三个随机氨基酸被加到 gp41 的抗原决定簇的前后，一并插入到血凝素的抗原性位点 B。通过直接重组法将含抗原决定簇的血凝素融合基因，连接到杆状病毒基因组内多角体基因启动子之下，获得一个三肽的随机文库。用 mAb2F5 抗体和抗人 lgG-FITC 偶联体处理被杆状病毒文库感染的昆虫细胞，荧光激活细胞分类仪筛选被病毒感染的细胞，只通过一轮筛选，就从 8000 个可能的变异体中得到了一个与抗体结合力明显增强的抗原决定簇，显示了杆状病毒应用于真核生物多肽文库的巨大潜力。随着重组杆状病毒技术的发展，限制杆状病毒用于多肽文库的转染效率和重组频率这两个瓶颈已经解决，用直接连接法进行重组杆状病毒的转染实验表明，1pg 病毒 DNA 可至少获得 2×10^5 个独立重组斑。

（3）多肽药物和疫苗研制　昆虫与哺乳动物在进化上分歧很远，基因组同源性很差，交叉反应少，对人不存在潜在的生物危害。杆状病毒表达系统自从问世以来，由于其不具有任何对哺乳动物有潜在致病性的动物病毒片段，得到了人们的高度重视。已有多个 BEVS 药物的疫苗进入Ⅱ期临床且规模相当大，BEVS 产品已接近产业化，前景乐观。

（4）有待解决的问题　杆状病毒是森林昆虫密度的主要调节因子，基因组相对较大，其中有许多基因是为了虫体或宿主特异性感染的，作为载体病毒来说反而是不利的，影响了外源的表达量；杆状病毒特有的两相感染模式和两种不同表型的病毒粒子均使得在特定的场合许多病毒基因是非必需的；杆状病毒感染细胞的晚期引起细胞裂解的病毒成分对于表达来说是有害的，目前已查明 *p10* 基因可能与感染后期的细胞裂解有关，有一缺失 *p10* 基因的突变病毒并不引起受感染细胞的裂解，分析发现 *p10* 的第 52～第 79 位氨基酸残基形成的结构域对细胞的解体起主要作用。在细胞工程实践中，连续高滴度感染细胞引起缺陷病毒的产生，通过对缺陷病毒粒子的基因组分析发现约 46%的基因组片段已经丢失，并存在着多个高丢失区段，这一切均提示我们尽可能删去一些与外源基因表达无关的甚至有害的片段，尤其是与受感染细胞裂解有关的部分，可在细胞或（宿主）中长期稳定表达外源基因的载体上构建，以提高 BEVS 的应用价值。

BEVS 的另一个不足之处是蛋白质的后加工与哺乳动物细胞不尽相同。一些主要的与糖基化有关的酶不是缺少就是活力很低，IE-1 控制的唾液酸糖苷转移酶重组到 AcNPV 中的 BEVS 已经成功地补救了上述缺陷。现在，BEVS 常用的启动子为多角体蛋白基因启动子，它虽然能超高效表达，但外源蛋白要在晚期能表达，此时受感染细胞正处于裂解死亡期，有必要寻找高效又能提前转录的启动子，来替代多角体蛋白启动子。表达系统的最终目的是为了获得纯的有生物活性的产品。用 6 个组氨酸残基、谷胱甘肽转移酶或麦芽糖结合蛋白作为

标签的融合蛋白表达载体已在 BEVS 中得到应用，但这些融合蛋白的产物很难直接用于临床试验。有必要发展和改进针对昆虫细胞和虫体的结构和功能蛋白与哺乳动物的异同，及虫体血淋巴中的高酚氧化酶活力等上述特性进行目的表达产物的纯化分离技术研究。

BEVS 的细胞表达相较于虫体表达有产物较易纯化、均一性好等优点，但昆虫细胞对血清的要求相对较高，胎牛血清的用量一般为 10%。在病毒感染复制阶段，降低胎牛血清的用量有助于提高重组病毒表达重组蛋白的水平，或野生型 AcNPV 的多角体蛋白。因此，开发成本低廉、适应性良好的无血清培养基是今后细胞培养技术的关键。另一值得研究的领域是选育病毒复制性、分泌性能良好，耐机械剪切且适应用海水、砂糖、酵母抽提物和水解乳蛋白等简单组分，且能高压消毒的无血清简单培养基的 MTCM-160 的昆虫细胞系。

（二）家蚕核型多角体病毒（BmNPV）表达系统

家蚕在我国的饲养已有五千年的历史，据考证最早饲养的目的并不在于取丝，而在于获得蚕蛹食用，用现代化检测手段检定了家蚕、幼虫、蛹、家蚕 NPV 及其多角体蛋白的致畸、致变、致癌等遗传毒理效应，认为无“三致效应”及急、慢性毒理效应；从蚕体及其血淋巴中未检测到对人有致病性的细菌、真菌、病毒和其他人畜共患病原，也不存在许多哺乳动物细胞系可能含有的潜在致癌转化元件，是一个安全可靠的遗传工程载体。

家蚕在我国的数千年饲养过程中，培育了上千个具有不同特性的家蚕品种，从中可以选出对不同基因有不同表达效率的优良品种。由于虫体细胞的种类丰富，外源蛋白的翻译后加工比单一细胞系更加完善，同时也不存在病毒累代表达的所谓“传代效应”。家蚕丝腺是一个天然的高效生物反应器，每天最高可合成 200mg 丝蛋白，在短短的 4～5d 内积累的丝蛋白量可占整个虫体蛋白质总量的 70%左右，这种高效的蛋白合成和加工能力是良好的表达外源基因的生物反应器。

家蚕遗传背景清楚、形体大、发育快，1ml 血淋巴细胞的表达量比 1ml 悬浮细胞的表达量高 10 倍左右；而且饲养容易、成本低廉，已丧失野外生存能力，环境扩散的可能性小；家蚕可作为昆虫营养代谢和激素调控的模式昆虫，且对其营养要求发育生长过程有了全面的了解，是目前唯一可用人工饲料进行全龄大规模无菌培养的昆虫。

家蚕核型多角体病毒-家蚕表达系统是理想的基因表达系统，至今每年有数百个基因利用该系统进行表达和功能分析。由杆状病毒 AcNPV 介导的家蚕定点基因整合技术已经建立，重组 AcNPV 介导的外源基因技术提供了位点特异的家蚕基因敲除手段，为研究昆虫基因的结构和功能、大量连续表达外源基因提供了可能。

我国有利用家蚕的悠久历史和广阔资源优势，因此我国 BEVS 的发展应该结合实际情况，重点在家蚕 BEVS，家蚕-BmNPV 表达系统更具优势，在发展重组多肽药物、疫苗及临床检测药盒等方面是一条可行的生物技术创新的途径。

第三节　哺乳动物细胞表达系统

在原核、酵母、昆虫细胞及哺乳动物细胞 4 大表达系统中，前三者尽管表达的产品种类已较多、表达量也很高，但因为表达产物的蛋白质翻译后加工过程或是不可能，或是不完

全，常常会影响产品的质量、生物活性、药代动力学行为、体内的稳定性以及免疫原性等。而哺乳动物细胞培养环境和蛋白自然表达环境相似，所以蛋白质易于保持其天然构象和翻译后的修饰。目前已经投放市场以及正在进行临床试验的治疗用蛋白质生物药物，70%来自哺乳动物细胞表达系统。在哺乳动物系统中表达目的蛋白对研究蛋白的功能也极其重要。

用哺乳动物细胞作宿主的表达体系具有以下主要优点：①哺乳动物细胞能识别和除去外源基因中的内含子，剪接加工成成熟的 mRNA，这是原核细胞办不到的；②哺乳动物细胞表达的蛋白质在翻译后被加工的机会较多（如糖基化），可提高产品正确构型的概率，有可能被改造成类人体源型，加工后的蛋白质免疫原性好，约为酵母型的 16～20 倍；③哺乳动物细胞易被重组 DNA 质粒转染，具有遗传稳定性和可重复性；④经转化的哺乳动物细胞可将表达的产物分泌到培养基中，其提纯工艺简单，成本低。鉴于上述优点，世界卫生组织（WHO）生物制品研究小组于 1986 年在日内瓦召开会议，建议世界各国采用长期培养的哺乳动物细胞株，制备各种生物制品。

一、哺乳动物细胞表达系统的构成

（一）宿主细胞

至今批准的由动物细胞表达的产品，其宿主细胞几乎都是 CHO 细胞。但是在实践中，人们发现它的一些不足，如它体内存在着一种内肽酶，可以使分泌的单链尿激酶原（prourokinase）分解成双链的尿激酶（urokinase）；此外其悬浮培养时的产量会降低，大规模培养常常需要采用微载体。人们已发现不同的宿主细胞表达产品的产量和质量（如糖基化类型）差别很大，而且不同的细胞对大规模培养和纯化工艺要求也不同，因此对宿主细胞的研究、选择和改造已越来越受到重视。近年来一些特征鉴定良好的细胞系，都已被作为宿主细胞试用于真核细胞的表达系统中。现将这些细胞的一般特性简述如下。

1. CHO-K1 细胞

1957 年由 Puck 从中国仓鼠卵巢中分离的一株上皮样细胞，目前被广泛应用于构建工程细胞的是其缺乏二氢叶酸还原酶（$dhfr^-$）的营养缺陷突变株，可以在氨甲蝶呤（methotrexate，MTX）压力下使外源基因的拷贝数扩增，得到较高水平表达。

2. MDCK 细胞

它是由 Madin 和 Darby 于 1958 年从成年雌性的西班牙长耳狗的肾脏分离获得的，是贴壁生长的上皮样细胞，已成功地在微载体上增殖。该细胞能支持多种病毒的增殖，已被用来生产兽用疫苗。

3. BHK-21 细胞

1961 年由 Macpherson 等从 5 只无性别的、生长仅 1d 的地鼠幼鼠的肾脏分离，现在广泛应用的是 1963 年采用单细胞分离技术经 13 次克隆的细胞。原始的细胞株是成纤维样的细胞，并且是贴壁依赖性的，但以后经无数次传代后细胞可悬浮生长。该细胞被广泛用于增殖病毒和表达重组蛋白质。

4. C127 细胞

它来自 RⅢ 小鼠乳腺肿瘤细胞，特别适用于带有牛乳头瘤病毒（bovine papilloma virus，BPV）载体的转染。当用 BPV-1 病毒载体转染后，细胞的生长形态可发生显著变化，因此转染成功的细胞可通过特有的转化形态加以识别。

5. Namalwa 细胞

它是从患有 Burkitt 淋巴瘤的同名病人体内获得的一株类淋巴母细胞，含有部分 EB 病毒基因，但不产生 EB 病毒，安全性好；它可以用无血清培养基在悬浮状态下高密度培养；能高效地表达外源基因。

6. Vero 细胞

该细胞是 Yasumura 等于 1962 年从正常的成年非洲绿猴肾分离获得的，是贴壁依赖性成纤维细胞，并可持续地进行细胞培养。它可支持多种病毒的增殖并制成疫苗，包括脊髓灰质炎病毒、狂犬病病毒和乙脑病毒等疫苗。

7. 鼠骨髓瘤细胞

选择骨髓瘤细胞作宿主细胞有以下优点：①它们是“专业”的分泌细胞，而且容易转染、生长，可在无血清培养基中高密度悬浮培养；②能对蛋白质进行糖基化修饰；③高效表达 Ig 基因的调控成分明确，有利于表达载体的增强子和启动子的合理设计。目前可供使用的骨髓瘤细胞有 Sp2/0、J558L 和 NS0 等。

8. COS 细胞

1981 年，Gluman 用复制起点缺失的 SV40 基因组 DNA 转化非洲绿猴肾细胞 CV-1，获得了 3 个细胞系：COS-1、COS-3 和 COS-7。它们被广泛地用于瞬时表达系统。原因是：①COS 细胞来源广、易培养、易转染；②COS 细胞能组成性表达 SV40 大 T 抗原，能使大多数哺乳动物细胞携带有 SV40 *ori* 复制子的质粒以附加体形式高拷贝扩增；③大量扩增的质粒及高表达的 mRNA 和蛋白质，很容易回收和分析。

（二）表达载体

为使哺乳动物细胞高水平地表达外源蛋白，首先要将其构建在一个高效表达载体内。它有赖于我们对各个调控元件的结构和作用机理的充分了解，以及将它们与外源基因合理地组建。经过几十年的努力，一系列高效表达的载体已经出现，表达外源蛋白的水平已从原来的“μg/L”级提高到了今天的 100mg/L 级，甚至“g/L”级水平，这使采用哺乳动物细胞生产生物药物的设想真正成为了现实。

哺乳动物细胞基因表达系统的发展与真核基因表达调控的研究紧密相关。表达载体的基因表达调控机理已研究得较清楚，构建的真核表达载体也不少，人们可以选择不同的载体、不同的增强子和启动子来获得外源基因的高效表达。以下对表达载体的构建组成及影响真核基因表达的因素作一简单介绍。

1. 表达载体的种类

（1）病毒载体（viral vector） 高等真核细胞中并不携带相当于细菌质粒的附加体 DNA，因此通常应用来自 DNA 病毒或 RNA 病毒基因作为载体。由于病毒在进化上获得了适应机体的完备机制，加之高转染效率和良好的靶向性使得通过改造后的病毒成为了一类很有价值的载体。目前常用的动物病毒表达载体大体上可分为两大类：一类是整合型（integrated），即外源基因通过这种病毒载体与宿主细胞的染色体整合，外源基因随宿主细胞基因组的复制而扩增，这类载体的代表是 pSV 系列；另一类为游离型（transient）或称病毒颗粒型，这类载体携带外源基因后，本质上仍然是一种完整或缺损的病毒，能够以病毒颗粒的形式在宿主细胞内自行复制或在辅助病毒存在下进行复制。常见的这类载体系统为痘苗病毒、腺病毒、杆状病毒及逆转录病毒等。

（2）质粒载体（plasmid vector） 哺乳动物细胞表达载体是从克隆载体上发展起来的，

因而含有原核基因序列，包括大肠杆菌的复制起始位点和便于在大肠杆菌中筛选的抗性基因（Amp^r）。另外，也含有真核细胞选择标记，如酶、抗生素等，大多数载体都带有 neo^r 基因，以便用于阳性细胞的筛选。有的载体还带有真核细胞复制的元件。因此，此种载体常常是穿梭载体（shuttle vector），即在细菌和哺乳动物细胞两者体内都能扩增。

2. 真核表达载体的基本成分

（1）一个真核表达载体首先必须含有基本的转录元件　①启动子和增强子。真核表达载体必须有真核启动子和增强子。原核启动子在哺乳动物细胞中是不起作用的。目前真核表达载体中常用的启动子和增强子有 SV40 病毒早期基因增强子（SV40）、Rouse 肉瘤病毒基因组长末端重复顺序（RSV）、人类巨细胞病毒（CMV）等。这些启动子和增强子是一些来源于病毒的启动子和增强子，宿主范围较广，在多种细胞中都可有一定的活性。②转录终止和加 poly（A）信号。真核基因在表达过程中，RNA 聚合酶通常跨过 poly（A）位点继续进行转录，确切的转录终止位点至今还不清楚。但成熟的 mRNA 的 3′ poly（A）尾是经过位点特异性的转录后切割并加上 poly（A）而形成的。poly（A）的形成有赖于 mRNA 中不同的序列，一是位于 poly（A）位点下游的 G U 富集区或 U 富集区，二是 poly（A）位点上游 11～30 个核苷酸处的一段高度保守的六核苷酸序列，即 AAUAAA，它被认为是加尾信号。真核表达载体必须带有 poly（A）位点下游序列，以保证新转录的 mRNA 能够有效地加上 poly（A）。另外，还需要剪接信号。

（2）允许载体在细菌体内扩增的质粒序列　包括能使质粒在细菌体内复制的起始位点和抗药性标记基因。常用于扩增的基因有二氢叶酸还原酶（*dhfr*）、谷氨酰胺合成酶和腺苷脱氨酶等基因。当采用 *dhfr* 基因时，可用氨甲蝶呤使之扩增；而该酶的扩增可防止 MTX 对细胞的毒害作用。在构建载体时可有意识地将 *dhfr* 基因与目的基因重组在一起。在培养基内逐渐增加 MTX 浓度，就会促使 *dhfr* 扩增，同时也会使其毗邻的目的基因随之扩增，从而大大提高了表达量。当采用谷氨酰胺合成酶基因时，可用硫胺甲硫氨酸（methionine sulphoxamine，MSX）使之扩增。此时还可带来另外两个好处，即可减少培养细胞对谷氨酰胺的需求；降低了氨对细胞的毒害。

（3）能用于筛选出外源基因已整合的选择标记　常用的筛选标记一般有两类。一类仅适用于密切相关的突变细胞株，如采用标记基因次黄嘌呤-鸟嘌呤磷酸糖基转移酶基因（*hgprt*）、胸苷激酶基因（*tk*）等基因，它们只分别适用于 $hgprt^-$、tk^- 等基因缺失的细胞株。另一类是显性作用基因，如新霉素磷酸转移酶基因（*neo*）基因，它能使新霉素磷酸化而失活。原来对新霉素敏感的哺乳动物细胞，一旦获得含该基因的载体后，就能在含该抗生素的培养基中存活。

（4）选择性增加拷贝数的扩增系统　有报道在载体上加入显性活动调控区（dominant control region，DCR）序列，可以克服因载体整合在宿主基因组不同位点而产生的表达水平降低。

pcDNA3.1 质粒是最常用的真核表达载体。该载体含有：一个高效表达的 CMV 启动子，用于目的基因插入的多克隆位点；原核细胞的复制起始位点；用于原核细胞筛选的抗性基因（Amp^r），用于真核细胞筛选的抗性基因（neo^r）以及小牛生长激素（BGH）的多聚腺苷酸化的位点（BGHpA）（图 7-10，来源于 Invitrogen 公司）。

3. 影响真核基因表达的因素

（1）转录元件对表达的影响　真核基因的启动子是在基因转录起始位点（+1）及其 5′

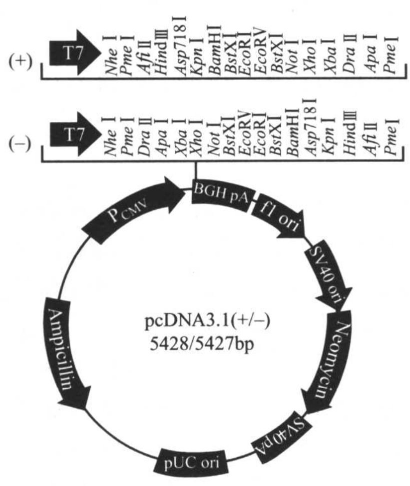


图 7-10　pcDNA3.1 质粒

上游附近大约 100～200bp 以内的一组具有独立功能的 DNA 序列，通常由核心启动子和上游启动子两部分组成，核心启动子是指足以使 RNA 聚合酶Ⅱ转录正常起始所必需的最少的 DNA 序列，其中包括“转录起始位点”。上游启动子元件决定 RNA 聚合酶Ⅱ的转录起始位点，通常包括位于 −70bp 附近的 CAAT 盒（CCAAT）和 GC 盒（GCGCGG）。当前，构建载体的许多启动子序列都来自病毒，包括逆转录病毒的长末端重复序列（long terminal repeat sequences，LTR）、SV40 病毒的早期和晚期启动子、腺病毒的晚期启动子、单纯疱疹病毒胸苷激酶（tk）启动子和人巨细胞病毒早期区启动子等。也有其他来源的启动子，如来自鼠细胞的金属硫蛋白基因启动子等。启动子的强弱决定了外源基因的表达水平。不同宿主细胞需要搭配不同种类的启动子，以充分发挥不同启动子的特异性和效率。

增强子可以在转录起始位点上游或离启动子相距很远的位置上起作用，它可使启动子的转录水平提高上千倍。由于它常常只能在特定的细胞和组织中起作用，因此与启动子一样，应根据不同的宿主细胞加以选择。

poly（A）序列可保持 mRNA 的稳定性，防止 mRNA 被降解。但过长的 poly（A）序列会过多消耗细胞的能量，不利于细胞转录的高效率。因此载体中 poly（A）序列长度必须适当。

终止序列要保证转录在正确位点停止，防止转录通读。

（2）剪接信号的利用和对表达的影响　真核基因表达调控中，初级转录物上加上 poly（A）以后，必须通过剪接除去内含子，才会产生成熟的 mRNA。许多真核表达载体中带有一个 SV40 的内含子，但因大多数插入的外源基因是 cDNA，所以一般也用不着剪接信号，但利用剪接信号这一特性构建的双顺反子（dicistronic）表达载体，筛选得到的阳性细胞株的表达水平比普通的 DHFR 双扩增载体要高，而且将表达基因与筛选基因连接在一起有助于减轻筛选强度。最稳固连接是将两个基因置于同一启动子下，内部核糖体进入位点（internal ribosomal entry site，IREs）使同一 mRNA 中的第二个基因得到翻译。该载体将 DHFR 基因与目的基因串联在一起，受同一个启动子控制，DHFR 的 cDNA 插在表达载体 pRK 的内含子中，两边为剪接供体 SD 和剪接受体 SA，这样，只有全长 mRNA 可以翻译出 DHFR，而剪接后的 mRNA 不翻译 DHFR，只翻译目的蛋白；被剪接的 mRNA 占全长 mRNA 的 95%，这样 DHFR 的表达减弱，只有那些含高拷贝的克隆可在 MTX 的选择压力下生存下来，经过随后的克隆富集之后，双顺反子载体中所携带的外源基因的表达量提高了 11～13 倍。这种多顺反子表达系统正被应用于多组分多亚基蛋白的偶联表达、将筛选标记置于最后一个顺反子中以利工程细胞的克隆、复杂调控环路的设计，如哺乳动物类细胞中表达基因的正反馈系统，以及使用正义或反义核酸及核酶技术对工程细胞进行多基因代谢工程等。

（3）翻译过程对表达的影响　真核生物的蛋白质合成遵循着滑动搜索模型，以真核 mRNA 为模板的翻译启动开始于最靠近其 5′端的第一个 AUG 密码，AUG 的旁侧序列与翻

译起始有密切关系。Kozak 通过对 699 种脊椎动物 mRNA 翻译起始密码 5′端及 3′端两侧核苷酸序列的分析，揭示了−3 位的 A 和+4 位的 G 对于 AUG 被起始识别有显著的促进作用，是保证外源基因有效转录的先决条件。

5′-非翻译区的二级结构对翻译起始也有影响，当 5′-UTR 中存在碱基配对时，会形成发夹式或茎环式二级结构，这类结构会阻止核糖体 40S 亚基的迁移，对翻译起始起顺式阻抑作用。因此，在构建表达载体时，需防止 5′-非翻译区出现二级结构。另外，第一个 AUG 至 5′端起点之间的长度（称为先导序列长度，leader length）同样影响起始效率和翻译起始的位点。因此，一个适当长度的、无高级结构的 AUG 密码子上游的 5′-非翻译区对于 mRNA 的有效翻译仍是必需的。另外，3′-非翻译区对相似核基因的翻译也有作用，尤其是 UA 序列，在许多编码细胞因子 mRNA 的 3′-UTR 中富含 UA 的保守序列，它常由几个相间分布的 UUAUUUAU 核苷酸序列组成，去除这段序列，可以明显提高 mRNA 的稳定性。我们不妨在表达一段外源基因时，只截取它的 ORF 区域，去除 mRNA 的 3′-UTR 序列，设计好起始密码 AUG 及其旁侧序列，以及终止密码 UGA、UAA、UAG 就可以了。当然，也有些外源基因的表达是受基因 3′-UTR 序列正调控的。

（三）目的基因

插入到真核表达载体的目的基因可以是真核基因，也可以是原核基因。真核基因也必须是 cDNA。无论是真核基因还是原核基因，都必须带有 5′端的非编码序列和 3′端的非编码序列。目的基因的两侧应带有与载体相同的酶切位点，以便与载体连接，形成重组体。

二、常用的细胞表达系统

1. CHO/DHFR 和 CHO/NEOSPLA 表达系统

当前，采用 CHO 细胞作为宿主细胞最为普遍，此时应用较多的表达系统是 CHO/DHFR 和 CHO/NEOSPLA 表达系统。第 1 个获准上市的生物药品 t-PA 就是用 CHO/DHFR 系统生产的。该系统采用 SV40 启动子控制外源基因的表达，*dhfr* 基因插入一个内含子中，载体转录成一个双顺反子结构，其中鼠 *dhfr* cDNA 5′是剪接供体，3′是剪接位点。这样全长 mRNA 翻译成鼠 DHFR，剪接加工后 mRNA 只产生 t-PA 分子。

CHO/NEOSPLA 系统是将低表达的新霉素磷酸转移酶（neomycin phosphotransferase，NEO）作为筛选标记，并通过有意损伤显性选择标记 *neo* 基因的表达，筛选高表达位点。损伤显性选择标记 *neo* 基因的表达的一种方法，是通过突变损伤 *neo* 翻译起始位点。该方法可获得 25～500mg/L 的表达水平。第二种方法是在 *neo* 编码序列中引入一个人工内含子，将有关基因和 *dhfr* 基因插入内含子内部，该策略可获得 2mg/ml 单抗表达水平。另有报道利用 NEOSPLA 表达载体，使 CHO 细胞表达单克隆抗体的产量达到了 1g/L。

2. C127/BPV-1 系统

该系统使用牛乳头瘤病毒-1（BPV-1）表达载体，用鼠金属硫蛋白启动子（mMT）启动外源基因的表达。在重组细胞中可检测到微弱的 BPV-1 早期基因转录，但无晚期基因转录，因此该系统是一个安全而合适的表达系统。用 BPV-1 构建的载体曾用于表达人的 β 干扰素、乙肝表面抗原和 t-PA、EPO 等。

3. 骨髓瘤细胞表达系统

骨髓瘤细胞高效表达 Ig 基因的调控成分明确，有利于表达载体的增强子和启动子的合

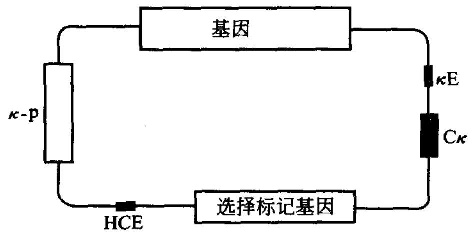


图 7-11　骨髓瘤细胞常用载体结构示意图
κ-p—Igκ 可变区基因启动子；HCE—Ig 重链增强子；κE，Cκ—κ 链稳定区 3′端序列

理设计。载体中采用 Ig 重链增强子，Ig κ 可变区基因启动子，在多克隆位点后有 κ 链稳定区 3′端序列，内含加 polyA 的信号和转录后 3′端正确加工的信号。Ig、RSV 和 CMV 启动子在骨髓瘤细胞中作用都很好（图 7-11）。另一个骨髓瘤细胞表达系统是 NS0/GS 表达系统。在该系统内，外源基因由 hCMV 5′序列和 SV40 3′序列调控，并有 SV40 晚期启动子调控筛选基因谷氨酰胺合成酶（GS）的表达。1992 年用该系统成功地表达了人-鼠杂合抗体和 rmAB-3 单抗，且有较高的表达量。

三、基因的导入方法

在真核细胞的表达研究中，细胞转染是一个关键步骤。在基因治疗中，也需要用基因转移技术导入外源基因。因此，该技术的研究已越来越受到重视，至今已开发了许多新的技术和方法。不同的细胞系，摄取和表达外源 DNA 的能力可相差几个数量级。因此，如果采用某种特定的细胞系，就一定要比较几种不同方法的效率。当前，用于 DNA 转染哺乳动物细胞的主要方法详见第六章。

四、外源基因在哺乳细胞的表达和基因表达产物的检测

外源基因在哺乳动物细胞的表达是一个多步骤、多环节的复杂过程，并不是每次实验都能得到理想的结果。一种克隆的外源基因是否能够在新宿主细胞中有效地表达，需要通过一定的实验才能作出正确的判断。如果人们事先已经知道克隆基因编码的蛋白质或 RNA 种类，就能够据此设计出适当的实验，检测基因在新宿主细胞中的表达情况。

1. 免疫荧光抗体法

此法方便易行，无需进行蛋白质的提取，适用于蛋白质的定性研究。

2. 免疫沉淀法

该法可用于检测并定量分析多种蛋白质混合物中的靶抗原。这种方法很敏感，可检测出 100pg 的放射性标记蛋白。当与 SDS-聚丙烯酰胺凝胶电泳并用时，即可用于分析外源基因在原核和真核宿主细胞的表达情况。

3. Westerm 印迹法

该法是在蛋白质凝胶电泳和固相免疫测定的基础上发展起来的。它结合了凝胶电泳分辨力高和固相免疫测定的特异敏感等多种优点，与免疫沉淀法比较，这种方法无需对靶蛋白进行同位素标记。Westerm 印迹法具有从混杂抗原中检测出特定抗原，或从多克隆抗体中检测出单克隆抗体的优越性，还可以对转移到固相膜上的蛋白质进行连续分析，具有蛋白质反应均一性、固相膜保存时间长等优点，因此该技术被广泛地用于蛋白质研究、基础医学和临床医学的研究。

五、哺乳动物细胞表达系统的改进策略

（一）改进表达载体，提高表达水平

为提高表达水平，主要要靠对表达系统的改进，包括采用更好的启动子和增强子，更好

地利用扩增系统或寻找高表达位点，以及采用更好的宿主细胞等。鼠骨髓细胞 MPG-11 产生 IgG 的能力为 5pg/（细胞·min），因此用 10L 反应器培养，当细胞密度为 10^7 个/ml 时，按理 1d 的产量应该达到 720g，但实际情况是常常连 1g 都达不到。这说明提高细胞的表达水平的潜力还是很大的。

1. 采用更强的启动子和增强子

作为表达载体，启动子既需要有强的转录活性，又要具有比较广的应用范围；除了病毒来源的强启动子外，现在人们热衷于寻找细胞内源性的强启动子，例如 H2-K6（小鼠主要复合物）启动子、鼠 pgK-1 启动子、鸡胞浆 β-肌动蛋白启动子等。真核表达载体 pCAGGS 使用的启动子是鸡的 β-肌动蛋白启动子 Achieve 与 CMV 的增强子元件组合在一起构建成的杂合启动子，它的转录活性是 CMV 启动子的 1.5 倍，比 CMV 启动子的应用范围广。鉴于 EF-1α 是真核细胞内含量最为丰富的蛋白质，在几乎所有的真核细胞中都有表达，pEF-Bos 使用人 EF-1α 启动子，使 gp130 和 GAM 在 COS 细胞中表达量明显高于其他真核载体。1996 年，Satoh 比较了不同增强子/启动子在 Namalwa 细胞表达 pro-UK 基因的效率。这些启动子是 SV40 早期启动子、人巨细胞病毒启动子、劳氏肉瘤病毒启动子、鸡 β-肌动蛋白启动子以及鼠白血病病毒启动子，结果发现鼠白血病病毒启动子介导的 pro-UK 表达水平是其他启动子的 10～20 倍，表达量高达 100μg/(10^6 个细胞·d)。其原因虽可能是在鼠白血病病毒启动子上有一些转录因子感应区，如在 LTR 中有 1 个甲状腺激素感应区（TRE）。当在培养基中加入甲状腺激素三碘甲状腺原氨酸时，能增加 pro-UK 的表达。

2. 更好地利用扩增系统或寻找高表达位点

外源基因的表达水平不仅与启动子的强度有关，还与基因的拷贝数有关。有两种方式可以提高载体的拷贝数，一种是采用与筛选标记基因共扩增的方式，常用的是 *dhfr* 基因，该系统是现今生物工程中应用得最多的，但它仅限于 $dhfr^-$ 细胞。另外，筛选抗性细胞费时费力，去除选择压力后，扩增基因常不稳定。另一种提高载体拷贝数的方法是采用自我复制型载体，载体以附加体形式在细胞内存在。如载体 pCEP4 中含有 EB 病毒的一个瞬时作用形式要素 *orip* 和其反式作用因子 EBNA-1 抗原，可以在各种哺乳动物细胞中以附加体的形式存在，拷贝数在 1～90 之间。此外，在 pCAGGS 的基础上加入 neo^r 筛选标志，得到的稳定表达载体 pABWN，载体中插有 BPV 的 DNA 片段（69%转化片段），该 DNA 片段使得质粒以 50～100 拷贝数在细胞中稳定存在，可有效地表达外源基因，并不会致细胞死亡；而 neo^r 筛选标记基因是在一个弱的启动子 TK 控制之下，加大 G418 的浓度，只有那些含高拷贝数的阳性细胞克隆可被筛选出来，人 IL-2 就在 pABWN 中得到了很高水平的表达。

将外源基因单一拷贝整合到染色体“热点”也可获得高水平的表达。这样就不需要或只需要少数几轮基因扩增，单拷贝或低拷贝外源基因即可获得高表达，减少了反复增强拷贝数的繁琐工作。筛选“热点”整合的新方法已经建立，即用传统方法转染和筛选高转录活性的染色体合适区域。载体含筛选基因和报告基因，两侧是人工重组位点如酵母的 F1p 重组靶点（FRT）或噬菌体 P1 LoxP 位点，它们能被重组酶 F1p 或 Cre 识别。重组位点之间的 DNA 可被第 2 个编码表达基因的载体精确置换，为避免两个 FRT 间插入序列被删除，两个 FRT 位点的核苷酸序列是不完全相同的（图 7-12）。

3. 改造现有的宿主细胞或采用理想的宿主细胞

采用理想的宿主细胞以及对现有的细胞进行改造也是提高表达水平的重要途径。用同样的载体在不同的细胞中表达尿型纤溶酶原激活剂（μ-PA），结果是 Sp2/0 最高，Namalwa

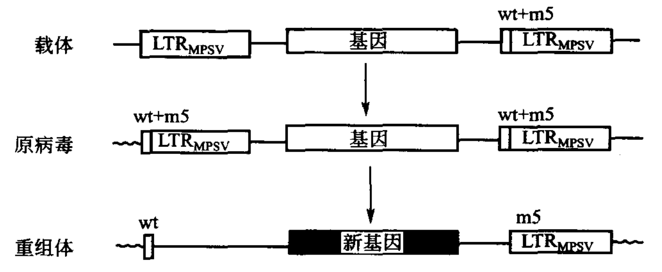


图 7-12　反转录病毒载体中的 FRT 序列及其原病毒形式
wt 是野生型 FRT 序列，m5 是 FRT 突变序列。染色体 DNA 用波状线表示
图中最上是载体形式、中间是包装和靶细胞中的原病毒形式、最下是重组交换后形式

次之，Vero 细胞最差。又如用 SV40 早期启动子在 Vero 细胞中表达 HBsAg，其表达水平仅 10ng/(10^6 个细胞·d)。当将这些细胞与非洲绿猴或狒狒的肝原代细胞融合后，表达水平提高了 50 倍，达到 514ng/(10^6 个细胞·d)。同样将骨髓瘤细胞 Sp2/O 与能表达 t-PA 的 CHO 细胞融合，也明显地提高了 t-PA 的表达量。

这些结果都表明，只要进一步改进表达载体的构建，选择理想的宿主细胞或改造现有的宿主细胞，哺乳动物细胞的表达水平就可能会有一个更加巨大的突破。

（二）改造宿主细胞及培养手段

哺乳动物细胞表达系统，不仅要在产品质量和安全性上具有竞争力，在经济性方面也需要有竞争力。除了提高外源基因的表达水平外，还需从改造工程细胞的代谢能力以及优化培养工艺着手，降低工程细胞大规模培养的成本。

1. 改造工程细胞

采用基因工程的手段，使工程细胞增加或减少某种酶，改造它的代谢能力和途径，使其降低对某些营养物质的需要，减少某些代谢产物的产生和危害，从而控制工程细胞的增殖速度和制造产品的能力，此即所谓“代谢工程”（metabolic engineering）。1996 年 Pak 等将胰岛素样生长因子和铁转递蛋白基因转入细胞，使之成为能自我分泌该两种因子的超级 CHO 细胞（super-CHO），从而可以将其直接培养于无血清培养基中，而无需另加这些因子。又如 1998 年 Cairo 等将谷氨酰转移酶基因转入杂交瘤细胞 KB-26.5，使其可在无谷氨酸的培养基中生长，且培养基中检测不出有害的氨。

通常，随着细胞生长密度的提高，产物也会随之增加；但是当超过一定限度时，不利的方面就会超过有利方面，如营养物质跟不上、有害代谢产物累积以致损伤细胞、细胞死亡增加、细胞内的蛋白水解酶影响到产品的质量等。因此，如何控制细胞的过度增长也是“代谢工程”的研究课题。干扰素调控因子 1（interferon regulatory factor 1，IRF-1）是一个可以抑制细胞生长的转录活化因子，将它与雌激素受体（estrogen receptor）基因构成融合蛋白后，在培养基中加入雌激素，就可诱导 IRF-1 的表达并抑制细胞的增殖，但不影响细胞合成和分泌所需产品的能力。另有报道，细胞周期蛋白依赖激酶（eyclin-dependent-kinase）p21、p27 及肿瘤抑制基因 *p53* 可使 CHO 细胞可逆地停滞于细胞周期中的 G1 期，可能因为它们的过量表达会抑制细胞周期蛋白 E-CDK2 复合体的磷酸化。当细胞过量表达 *p27* 基因产物后，细胞停止生长，但无细胞凋亡现象，此时所需产物的比生产率提高了 15 倍。另一控制细胞增殖的策略是利用 p27-bcl-XL-3 顺反子表达系统，转染该系统的 CHO 细胞的产物表达

水平比对照细胞高 30 倍。

尽管与微生物宿主细胞相比，通过遗传改造哺乳动物细胞“代谢工程”的研究还处于初级阶段，但上述成功的事例，说明这方面潜力还很大。

2. 改进培养工艺，降低生产成本

另一个降低成本的重要途径是优化培养工艺。例如，有人采用灌流式生物反应器，通过透析膜补充缺乏的营养成分，如葡萄糖、半胱氨酸和丝氨酸等，通过微硅胶管补充溶氧，实现了 Namalwa 细胞的高密度培养，使细胞密度保持在 10^7 个细胞/ml 以上，G-CSF 表达量从 $25cm^2$ 培养瓶中的 1.8mg/L 增加到 41mg/L。

3. 抑制细胞凋亡，延长培养周期

细胞凋亡（apoptosis）是一种由遗传基因决定的程序性细胞主动死亡。虽然它对多细胞生物的生命极其重要，但在生物工程中往往使细胞快速死亡而降低了生产效率。生物反应器中大多数细胞的死亡都是因细胞凋亡引起的。为防止细胞培养过程中的细胞凋亡，一般可采用如下三种重要措施：①通过培养基和氧的优化供应防止营养物质和氧的缺乏；②用化学添加剂如抗氧化剂等阻断细胞凋亡过程；③采用抗细胞凋亡基因改造工程细胞。抗细胞凋亡的原癌基因 *bcl*-2（B 淋巴瘤/白血病蛋白-2）被研究得最多，它在细胞凋亡级联反应中处于重要环节。通过遗传工程使哺乳动物细胞过表达抗凋亡基因 *bcl*-2 基因，虽不能防止细胞死亡，但能延长细胞寿命，增加产物产量。*bcl*-2 家族中的另外一些成员可能也有与 *bcl*-2 一样甚至更强的抑制细胞凋亡的作用。如 *bcl*-XL 在许多哺乳动物细胞中得到表达后，可对有些诱导因素如营养物缺乏、辐射和糖皮质激素等引起的细胞凋亡有抑制作用。另外有实验表明，当共表达 *bag*-1 和 *bcl*-2 基因时，可产生有相同的或协同的抗细胞凋亡作用。还有一些已被确认有抑制细胞凋亡的基因包括 Dad1、A20 和 SPI-1 等。

（三）提高产品质量

改进宿主细胞对外源表达产物的加工、修饰过程，使其尽可能具有与天然产物相似的理化性质和生物学活性，可提高产品的质量。在此过程中，糖基化有无以及类型对许多糖蛋白类生物药物来说，意义更大。蛋白质有两种糖基化方式：*N*-糖基化和 *O*-糖基化。一般地说，*O*-糖链对蛋白质的特性影响不大，而 *N*-糖链的不同对产品可产生较大的影响。决定糖链类型不同的主要因素有三个：①合成肽链的不同；②细胞内糖基化酶的不同；③细胞培养环境的影响。

在现有的哺乳动物表达系统中，用 CHO 细胞表达的糖蛋白，因其缺少 α-2,6-唾液酸转移酶的功能，其类型与人不尽相同。而 Namalwa 细胞表达的 t-PA、pro-UK 和 EPO，它们的糖基化均和人的自然产品一致。此外，人们正打算采用“糖基化工程”（glucosylation engineering），即用基因工程的手段，人为地改变肽链结构、增加某些酶基因，以及改进和控制某些培养条件等，以达到正确糖基化的目的。如在 t-PA 基因中进行点突变，改变了一个氨基酸，使之增加一个糖基化位点，此时的 t-PA 在血浆中的清除率，较之原先的 t-PA 减少了 10 多倍。而将 t-PA 与 α-2,6-唾液酸转移酶基因共转染 CHO 细胞，表达的 t-PA 的糖基化情况与人类接近。另外，控制和改变细胞的培养环境对 *N*-糖基化也可产生一定影响。如通过添加地塞米松、改变培养基糖的成分、使培养温度突然下降等，都可使 pro-UK 的糖基化形式从含岩藻糖的 2 分支复杂型寡糖链，转变为含岩藻糖的 3 或 4 分支复杂型寡糖链，从而使生产的重组蛋白质糖基化形式与人类的一致。

（四）筛选高表达细胞克隆的新方法

近来有如下一些新方法：采用绿色荧光蛋白（green fluorescence protein，GFP）分离方法；在细胞克隆上铺一薄层琼脂糖，用膜转移使表达产物吸附，再用免疫学方法检出高表达细胞克隆株；有的用特异抗体吸附于细胞膜，再加入二抗或磁珠来筛选高表达的单个细胞；有的用二抗标记的凝胶小滴包裹单个细胞，再用流式细胞仪（flow cytometry）筛选高表达细胞等。

第四节　重组蛋白的检测与鉴定

在进行外源基因表达产物的分离纯化之前，必须明确宿主细胞是否表达目的蛋白质，表达量如何；在分离纯化过程中，必须明确每个步骤的目的蛋白质的浓度、纯度和生物活性等；最终纯化重组蛋白质在浓度、纯度和生物活性满足要求后，还必须通过肽谱分析、氨基酸序列分析、分子量和等电点测定等方法鉴定得到的蛋白质就是我们需要的目的蛋白质。

一、定量分析

蛋白质的定量分析是生物化学和其他生命学科最常涉及的分析内容，是临床上诊断疾病及检查康复情况的重要指标，也是许多生物制品，药物、食品质量检测的重要指标。在生化实验中，对样品中的蛋白质进行准确可靠的定量分析，则是经常进行的一项非常重要的工作。

蛋白质是一种十分重要的生物大分子，它的种类很多，结构不均一，分子量又相差很大，功能各异，这样就给建立一个理想而又通用的蛋白质定量分析方法带来了许多具体的困难。目前测定蛋白质含量的方法有很多种，如紫外分光光度法、凯氏定氮法、双缩脲法、Lowry 法、BCA 法、胶体金法、考马斯亮蓝染色法、银染法、荧光法、电泳法等。但是每一种方法都有其特点和局限性，因而需要在了解各种方法的基础上根据不同情况选用恰当的方法，以满足不同的要求。

通过定量分析可以分析鉴定外源基因是在胞外、胞内或细胞周质表达，以及表达量的高低，并可以测定每个分离纯化步骤以及最后纯化的目的蛋白质的含量和浓度，这对于整个分离纯化过程的设计和质量控制是非常重要的。常见的定量分析方法参见第一章相关内容。

二、纯度分析

重组蛋白质的纯度是一个重要的指标，也是一个相对的指标。纯度分析通常采用还原性或非还原性 SDS-PAGE、PAGE、等电点聚焦（IEF）、毛细管电泳（CE）和各种 HPLC 等。

蛋白质纯度通常用百分数表示，其准确性有时与检测方法、定量方法、所选用仪器的参数有关。例如，用 HPLC 来分析蛋白质纯度时，一般在其吸收高峰 280nm 处检测，但这很可能不是非蛋白质杂质甚至小分子物质的紫外吸收峰，这样就检测不到这些物质的存在，检测得到的蛋白质纯度就偏高。此外，在层析中普遍采用积分面积的方法，而仪器的参数设置

与所得的报告很有关系，设置不妥往往导致结果不够准确。

近年来，人们开始采用一些新的方法进行蛋白质纯度分析，如质谱等。特别是电喷雾电离质谱（ESI-MS）的发展，只要 pmol 级的样品就能测定。此外，测定蛋白质 N 端或 C 端是否均一也是很有效的方法。只用一种方法鉴定蛋白质纯度是不够的，至少应用两种或有两种以上不同分离机理的分析方法相互佐证，结果才比较可靠。

如果纯化的重组蛋白质用作治疗药物，仅仅测定蛋白质纯度是不够的，还必须对杂质进行检测和分析。杂质主要包括蛋白质和非蛋白质杂质。表 7-1 列出了通常需要检测的杂质及其检测方法。

表 7-1　外源基因表达产物中常见杂质和污染物及其检测方法

杂质和污染物	检测方法
蛋白质类杂质(如宿主细胞蛋白质和其他蛋白质类杂质	SDS-PAGE,HPLC,毛细管电泳(CE),免疫分析(ELISA 等)
内毒素)	鲎试剂、家兔热原法
残余 DNA	DNA 斑点杂交
蛋白质降解产物	SDS-PAGE,等电点聚焦(IEF),HPLC,N- 或 C-端氨基酸分析
氨基酸突变	肽谱、HPLC、CE、氨基酸分析、Edman 分析、质谱
产物变性或聚合脱氨基	SDS-PAGE、凝胶过滤层析、IEF、HPLC、CE、Edman 分析、质谱
甲酰基甲硫氨酸	肽谱、HPLC、IEF、CE
甲硫氨酸氧化	肽谱、氨基酸分析、HPLC、质谱、Edman 分析
单克隆抗体(亲和配基脱落)	SDS-PAGE、免疫分析
微生物污染	无菌试验
支原体	微生物学检查
病毒	微生物学检查

蛋白质类杂质主要是纯化过程中残余的宿主细胞蛋白质。此外，重组蛋白质本身也可能发生某些变化，形成在理化性质上与原蛋白质极为相似的蛋白质杂质，如由污染的蛋白酶所造成的产物降解、冷冻过程中由于脱盐而导致的蛋白质沉淀、冻干过程中处理不当所引发的蛋白质聚合等。这些由于降解、聚合或错误折叠而造成的蛋白质变构体在体内往往会导致抗体的产生，因此这类杂质也都得严格控制。它们的测定一般采用免疫分析、SDS-PAGE 和 HPLC 等方法。

非蛋白类杂质主要包括具有生物学作用的病毒和细菌等微生物、热原质、内毒素、致敏原及宿主细胞 DNA 等。可通过微生物学方法来检测生物工程药物中是否有外源病毒和细菌等污染；热原质可用传统的注射家兔法进行检测；鲎试验法可以测定内毒素；宿主细胞的残余 DNA 可以采用 DNA 斑点杂交法检测。

三、重组蛋白质的鉴定

主要通过对重组蛋白质的分子量、等电点以及氨基酸组成和序列分析等进行鉴定。

1. 分子量测定

蛋白质分子量测定最常用的方法有 SDS-PAGE 和层析方法（常规层析和 HPLC 中的凝胶过滤）。SDS-PAGE 方法测定蛋白质亚基的分子量，而凝胶过滤层析则是测定完整的蛋白质分子量。同时用这两种方法测定同一蛋白质的分子量，可以方便地判断样品蛋白质是否是寡聚蛋白质。此外，也可以用毛细管电泳测定蛋白质分子量，其准确性比 SDS-PAGE 好。目前最精确的测定方法是质谱。电喷雾电离质谱（ESI-MS）和基质辅助激光解吸离子化-飞行时间质谱（MALDI-TOF-MS）可以测定分子量 1 万～30 万的蛋白质，用量少（pmol 至

fmol 级），方法简单快速（只需数分钟），而且结果准确（相对误差<0.01%）。

2. 等电点测定

应用等电聚焦方法既可以测定蛋白质的等电点，同时还可以检测蛋白质的纯度。

3. 氨基酸成分分析

在氨基酸成分分析中，一般含 50 个左右氨基酸残基的蛋白质的定量分析接近理论值，即与序列分析结果一致；而含 100 个左右的氨基酸残基的蛋白质的成分分析与理论值会产生较大的偏差。分子量越大，偏差越严重，主要原因是不同氨基酸的肽键在水解条件下，有些水解不完全，有些则被破坏，很难做出合适的校正。但氨基酸成分分析对目的产物的纯度仍可以提供重要信息。完整的氨基酸成分分析结果，应包括甲硫氨酸、胱氨酸和色氨酸的准确值。氨基酸成分分析结果应为 3 次水解样品测定后的平均值。

4. 肽谱分析

肽谱是指用酶法或化学法降解蛋白质后，对形成的肽段进行分离分析的结果。它是检测蛋白质一级结构中细微变化的最有效方法。现在，肽谱主要应用 HPLC 和 CE 来测定；一般采用 RP-HPLC 与质谱联用的方法，RP-HPLC 根据肽段的长短和疏水性质来分离，收集各个肽峰进行质谱分析；但亲水性的肽段或疏水性很强的肽段用 HPLC 不易被分离，而 CE 有时可以弥补这个缺欠。

5. 部分氨基酸序列分析

部分氨基酸序列分析（如 N 端 15 个氨基酸残基以及 C 端若干个）可作为重组蛋白质的重要鉴定指标。

6. 质谱分析

如前所述，质谱可以精确地测定蛋白质的分子量。因此，若重组蛋白质发生缺失或替代等突变，可用质谱方法检测。此外，质谱还广泛用于蛋白质的纯度鉴定、分子量测定、序列测定、肽谱分析和二硫键分析，以及蛋白质乙酰化、糖基化、磷酸化分析等。

7. 蛋白质二硫键分析

二硫键和巯基与蛋白质的生物活性密切相关，重组蛋白的二硫键是否正确配对，对其生物活性起到至关重要的作用。测定巯基的方法有 5,5′-二硫-2-硝基苯甲酸法（DTNB）和 *N*-乙基顺丁烯二酰亚胺（NEMI）等。

（孙　军）

第八章　分子标记技术

近年来，随着分子生物学、细胞生物学、基础免疫学和免疫化学等学科的进展以及应用现代高新技术建立的仪器分析日趋发展，标记技术也不断完善和更新。各种新的标记物与标记方法不断涌现，至今已发展成为一类检测微量和超微量生物分子的重要技术，在分子生物学和其他生物学科的研究中得到广泛应用。

第一节　概　述

生物分子的标记方法大致分为两类：同位素标记与非同位素标记。同位素标记属于传统的标记方法，尽管目前仍在广泛应用，但由于同位素，特别是放射性同位素对人体的伤害较大，逐渐被其他安全高效的非同位素标记方法所取代。这已成为标记技术发展的必然趋势。非同位素标记方法中，各种发光技术潜力最大，发展也最为迅速。这些发光标记方法可细分为化学发光标记、生物发光标记、荧光标记等类别。

一、放射性同位素

同位素分为放射性同位素和稳定性同位素两种形式。放射性同位素（radioisotope）如 ^{3}H、^{14}C，经历着自身的衰变过程，并放射出辐射能，是不稳定的，具有物理半衰期，广泛用于各种体外检测技术中。稳定性同位素（stable isotope）如 ^{2}H、^{13}C、^{15}N 和 ^{18}O，以一定比例存在于自然界；因其无放射性，物理性质稳定，对人体无害，故常用于体内代谢过程的研究。

（一）放射性同位素的基本性质

放射性同位素的原子核很不稳定，会不间断地、自发地放射出射线，直至变成另一种稳定同位素，这就是所谓“核衰变”。放射性同位素在核衰变的时候，可放射出 α 射线、β 射线、γ 射线和电子俘获等，但是放射性同位素在进行核衰变的时候并不一定能同时放射出这几种射线。核衰变的速度不受温度、压力、电磁场等外界条件的影响，也不受元素所处状态的影响，只和时间有关。放射性同位素衰变的快慢，通常用“半衰期”来表示。半衰期（half-life）即一定数量放射性同位素原子数目减少到其初始值一半时所需要的时间。如 ^{32}P 的半衰期是 14.3d，就是说，假使原来有 100 万个 ^{32}P 原子，经过 14.3d 后，只剩下 50 万个了。半衰期越长，说明衰变得越慢，半衰期越短，说明衰变得越快。半衰期是放射性同位素的一特征常数，不同的放射性同位素有不同的半衰期，衰变的时候放射出射线的种类和数量也不同。分子标记技术中常用的放射性同位素的半衰期及其他特征见表 8-1。

表 8-1 常用同位素的特征

同位素	半衰期	β射线能量/MeV	同位素	半衰期	β射线能量/MeV
^{3}H	12.3年	0.018	^{35}S	87.1d	0.167
^{14}C	5720年	0.156	^{131}I	8.05d	0.605
^{32}P	14.3d	1.71			

（二）放射性强度及其度量单位

放射性同位素原子数目的减少服从指数规律。随着时间的增加，放射性原子的数目按几何级数减少，用公式表示为 $N=N_0e^{-\lambda t}$，这里，N 为经过 t 时间衰变后，剩下的放射性原子数目，N_0 为初始的放射性原子数目，λ 为衰变常数，是与该种放射性同位素性质有关的常数，$\lambda=y(t)=e^{-0.693t/\tau}$，其中 τ 指半衰期。放射性同位素不断地衰变，它在单位时间内发生衰变的原子数目叫做放射性强度（radioactivity），放射性强度的常用单位是居里（Curie），表示在 1s 内发生 3.7×10^{10} 次核衰变，符号为 Ci。$1Ci=3.7\times10^{10}dps=2.22\times10^{12}dpm$，$1mCi=3.7\times10^{7}dps=2.22\times10^{9}dpm$，$1\mu Ci=3.7\times10^{4}dps=2.22\times10^{6}dpm$。1977 年国际放射防护委员会（ICRP）发表的第 26 号出版物中，根据国际辐射单位与测量委员会（ICRU）的建议，对放射性强度等计算单位采用了国际单位制（SI），我国于 1986 年正式执行。在 SI 中，放射性强度单位用贝可勒尔（Becquerel）表示，简称贝可，为 1s 内发生一次核衰变，符号为 Bq。$1Bq=1dps=2.703\times10^{-11}Ci$。该单位在实际应用中减少了换算步骤，方便了使用。

（三）射线与物质的相互作用

放射性同位素放射出的射线碰到各种物质的时候，会产生各种效应，它包括射线对物质的作用和物质对射线的作用两个相互联系的方面。例如，射线能够使照相底片和核子乳胶感光；使一些物质产生荧光；可穿透一定厚度的物质，在穿透物质的过程中，能被物质吸收一部分，或者是散射一部分，还可能使一些物质的分子发生电离；另外，当射线照射到人、动物和植物体时，会使生物体发生生理变化。射线与物质的相互作用，对核射线来说，它是一种能量传递和能量损耗过程，对受照射物质来说，它是一种对外来能量的物理性反应和吸收过程。

各种射线由于其本身的性质不同，与物质的相互作用各有特点。这种特点还常与物质的密度和原子序数有关。α 射线通过物质时，主要是通过电离和激发把它的辐射能量转移给物质，其射程很短，一个 1 兆电子伏（1MeV）的 α 射线，在空气中的射程约＜1.0cm，在铅金属中只有 23μm，一张普通纸就能将 α 射线完全挡住，但 α 射线的能量能被组织和器官全部吸收。β 射线也能引起物质电离和激发，与 α 射线的能量相同的 β 射线，在同一物质中的射程比 α 射线要长得多，如＞1MeV 的 β 射线，在空气中的射程可达数米，高能量快速运动的 β 粒子，如 ^{32}P，能量为 1.71MeV 遇到物质，特别是突然被原子序数高的物质（如铅，原子序数为 82）阻止后，运动方向会发生改变，产生韧致辐射。韧致辐射是一种连续的电磁辐射，它发生的概率与 β 射线的能量和物质的原子序数成正比，因此在防护上采用低密度材料，以减少韧致辐射。β 射线能被不太厚的铝层等吸收。γ 射线的穿透力最强，射程最大，1MeV 的 γ 射线在空气中的射程约有几十米之远，γ 射线作用于物质可产生光电效应、康普顿效应和电子对效应，它不会被物质完全吸收，只会随着物质厚度的增加而逐渐减弱。

（四）射线的防护

放射性物质可以从体外进入体内，放出射线，对人体造成损害。就外照射而言，由于各

种射线穿透能力不同，γ射线照射对机体的危害大于β射线，而β射线的危害性又大于α射线。受照射部位不同，受害程度也不同，对某种放射性同位素蓄积率高的组织或器官，必然受害严重，如^{32}P对骨骼系统危害较大，^{125}I和^{131}I主要危及甲状腺器官等。但是，由于射线与机体作用可产生电离，同样具有较大的危害，当放射性物质进入了体内，对机体造成内照射的情形下，α射线由于电离能力强，其危害性大于β射线和γ射线，而β射线的内照射危害又大于γ射线。放射防护除了保护操作者本人免受辐射损伤外，还要防止射线对周围人群的伤害及对环境的破坏。要注意做好放射性污物、污水的收集与处理，避免环境污染。避免不必要的射线照射，减少人群的剂量负担。

国际放射防护委员会 1977 年第 26 号出版物中提出防护的基本原则是放射实践的正当化、放射防护的最优化和个人剂量限制。这三项原则构成剂量限制体系。

1. 放射实践的正当化

在进行任何放射性工作时，都应当进行代价和利益的分析，要求任何放射实践，对人群和环境可能产生的危害比起个人和社会从中获得的利益应当是很小的，即效益明显大于付出的全部代价时，所进行的放射性工作就是正当的，是值得进行的。

2. 放射防护的最优化

使放射性和照射量在可以合理达到的尽可能低的水平，避免一些不必要的照射，要求对放射实践选择防护水平时，必须在由放射实践带来的利益与所付出和健康损害的代价之间权衡利弊，以期用最小的代价获取最大的净利益。最优化原则又称为 ALARA 原则，其目的在于促进社会公众集体安全的卫生保健，它是剂量限制体系中的一项重要的原则。

3. 个人剂量限制

个人剂量限制是强制性的，必须严格遵守。ICRP 规定工作人员全身均匀照射的年剂量当量限制为 50 毫希沃特（mSv），居民的年剂量当量限值为 1mSv。我国放射卫生防护基本标准中，对工作人员年剂量当量限值，采用了 ICRP 推荐规定的限值，为防止随机效应，规定放射性工作人员受到全身均匀照射时的年剂量当量不应超过 50mSv，公众中个人受照射的年剂量当量应低于 5mSv。当长期持续受放射性照射时，公众中个人在一生中每年全身受照射的年剂量当量限值不应高于 1mSv，且以上这些限制不包括天然本底照射和医疗照射。

二、放射性同位素标记

在分子生物学研究中，各种生物分子的检测分析以及动态观测常用各种探针或示踪物，而对这些探针或示踪物的标记最早采用的方法就是放射性同位素标记。探针（probe）是一种已知的特异性的分子，它带有合适的标记物以检测与其特异性相互作用的待测分子（靶分子）。探针和靶的相互反应包括抗原-抗体、血凝素-碳水化合物、亲合素-生物素、受体和配体，以及核酸与其互补核酸间的杂交等反应。由于探针的广泛应用及在生物学领域的重要研究价值，致使探针的标记技术成为分子生物学领域重要的基本技术。

放射性同位素多用于核酸探针的标记，常对核酸合成所需的原料 dNTP 进行标记，如^{32}P-dNTP、^{3}H-dNTP、^{35}S-dNTP 等。用放射性同位素标记核酸有以下优点。

（1）灵敏性高　一般可达到 0.5～5pg 或更低剂量核酸的检测水平，可以检测极少量或拷贝数少的基因组（可延长曝光时间或使用增敏屏增敏）。

（2）特异性高　用放射自显影法，样品中存在的无关核酸或非核酸成分不会干扰检测结果，准确率高，假阳性率低。

（3）方法简便　各种同位素标记物已经商品化，随时购买，使用方便；各种检测方法也相对较为简单。

所以，放射性同位素标记核酸探针在一些有条件的单位，作为主要的标记方法仍在使用。但是，放射性同位素标记技术存在一些重要的缺点。①半衰期短，必须经常标记探针。常用的^{32}P半衰期只有14.3d，^{35}S的半衰期尽管长达88d，但其衰变能量只有^{32}P的1/10，灵敏度较低，应用范围狭窄，只能用于多拷贝基因的检测。^{3}H也是如此，其半衰期虽长达12.26年，但衰变能低，灵敏度太低。②检测时间长。用放射自显影需要较长的曝光时间（1～15d)。③危害较大。放射性同位素对人体有害，实验室和环境易被污染，放射性废物处理困难。因此，推广使用受到限制。

三、非同位素标记

由于放射性同位素标记技术的重要缺陷，非同位素标记技术的研制引起了高度重视并陆续取得进展。Langer等（1981）合成了生物素化dUTP和NTP类似物，开创了非同位素标记的新领域。用生物素标记核酸探针检测和识别特定的核酸序列，已广泛应用于临床检验和科学研究。Leary等（1983）用切口平移法标记的生物素DNA探针成功地用于Southern印迹杂交技术。Forst等（1985）发表了光敏生物素直接标记核酸的简便方法，使核酸和蛋白质的生物素标记探针技术前进一大步。Renz等（1984)、Thorpe等（1985)、Pollars和Knight等（1990）用化学法直接把辣根过氧化物酶接到DNA上，运用发光显影法提高了敏感性。Muhlegger等（1990）用地高辛（digoxigenin）标记DNA，再用地高辛抗体连接碱性磷酸酶，催化发光底物金刚烷脱磷并发光，灵敏度与^{32}P标记探针相同，而且^{32}P标记的DNA探针杂交膜需要在－70℃件下曝光4～7d于X射线胶片上显影，碱性磷酸酶发光法只需10～30min，十分快速、简便。

上述各种非同位素标记方法中，均是以各种发光材料作标记物，因此发光材料的发展成为非同位素标记技术发展的重要基础。从开始的几种有限的发光材料，如荧光素（fluorescein）和得克萨斯红（Texas red），发展到目前数量众多的各类发光材料，并建立了化学发光（chemiluminescent）标记与检测技术、生物发光（bioluminescent）标记与检测技术、荧光（fluorescent）标记与检测技术，较为著名的有安玛西亚（Amersham Pharmacia）公司开发的各种化学发光或电化学发光系统；PE公司与Pierce化学公司的各种基于溶液或斑点分析的高灵敏度检测系统；Promega与Pharmingen公司生产的荧光素酶（luciferase）生物发光系统；CLONTECH公司的绿/红荧光蛋白（green/red fluorescent protein）检测系统等。并且，随着研究的发展，许多生物技术公司将这些发光检测系统开发成各种面向用户的易于使用的标记试剂盒，使用更加方便、简洁，也使发光标记技术的应用范围大为大拓展，以致将传统的同位素标记世袭领域，如DNA序列测定等，逐渐纳入发光技术应用的范围。

（一）发光信号的特点

发光信号的检测涉及两种类型的技术——化学发光和荧光。在化学发光检测中，通过某种酶促反应，催化一种化学底物产生能够检测的光信号；而荧光检测中，则由仪器提供外部光源，如白炽灯或激光。当入射光被染料分子吸收后，自然地转变为另一不同波长的光，从而被检测或测量。

这两种检测方法各有各的特色。化学发光在检测的灵敏性和较宽的动力学范围方面更有优势，并且不用担心过多的背景光，因为化学发光来自于内部，尽管它没有荧光明亮，但它

的背景也低，且不会像荧光那样对样品造成损害。然而，荧光的检测与成像则较化学发光简单，应用也更为广泛，而且通过标记不同的荧光物质，还可以同时检测一个样品中的多种参数或指标，例如在一个样品中同时测量活细胞和死细胞时，可以通过标记一种红色的和另一种绿色的荧光染料就可以实现。综合来看，这两种标记方法互有利弊。随着研究的深入，化学发光物的数量开始逐渐增多，实验中选择哪一种技术基本上取决于检测的信噪比（ratio of signal to noise）。

（二）化学发光的本质

化学发光是指伴随化学反应过程所产生的光的发射现象。某些物质在进行化学反应时，吸收了反应过程中所产生的化学能，使反应产物分子激发到电子激发态。当电子从激发态的最低振动能级回到基态的各个振动能级时产生辐射，多余的能量以光子的形式释放出来，这一现象称为化学发光。

在一个典型的化学发光反应中，关键的化学试剂含有丰富的能量，并能以光的形式释放这些能量。这类化学试剂在酶作用之前是稳定存在的。酶催化这些物质的最终降解，并伴随着光的散发。这一简单的反应被应用到各种分析检测方法中，并随着技术的改进使检测更为灵敏。例如，Tropix公司从最初的单一发光底物发展到目前使用的高通量筛选分析检测系统，广泛用于报告基因评价、第二信使检测、免疫检测以及核酸检测分析等；Pierce Chemical公司也发展了多种化学发光检测试剂盒，用于 Western 斑点杂交、Northern 斑点杂交、Southern 斑点杂交检测以及各种液相检测。

高通量筛选标记是发光标记的一个重要发展方向，目前已有公司生产了用于微量平板的高通量标记方法，如针对异源 cAMP 分析的标记试剂以及可用于 G-蛋白偶联受体筛查检测未知配体标记方法。

（三）生物发光

生物发光（bioluminescent）是化学发光的一个方面，是指在生物体内由于酶作用于化学底物而产生的发光现象，如萤火虫的发光。化学发光只能在体外进行，而生物发光则发生在生物体内。以往化学发光和生物发光被看作是彼此独立的过程，随着这一领域研究的深入，逐渐加深了对各种生物发光反应机理的认识。实际上，生物体内的各种发光现象都与化学反应有关，本质上生物发光是发生在生物体内的一种化学发光。生物发光系统在研究基因表达方面有重要价值，它使传统的报告基因分析应用范围拓宽而覆盖至细胞生理学的广泛领域。

荧光素酶检测系统是生物发光检测系统中较为突出的一种简单、高活性方法。它灵敏度高，线性动力学范围超过 8 个数量级。而且，因为发光反应分两步完成，当酶不存在时，不会产生可检测的荧光，这导致了它极高的灵敏性。Promega 公司开发的荧光素酶检测系统具有代表性。它利用两种荧光素酶，一种来源于萤火虫，另一种来源于水母样海洋生物 Renilla。在报告基因评价中，可先加入萤火虫底物产生荧光，然后加入另一试剂淬灭萤火虫反应，并激发 Renilla 的荧光反应用于检测。由于荧光素酶发光反应非常迅速，因此利用荧光素酶的检测得到广泛的应用。

（四）荧光的基础

尽管化学发光与生物发光系统具有巨大的应用潜力，但目前广泛使用的仍然是荧光标记与检测系统，这与它们的发光方式有关。荧光的发光能力强并且是一个连续再生的过程，只要样品存在，当样品受到激光照射时，就会产生荧光；而化学发光只会激发一次。这样，就

使得荧光成像较为容易，并能够在细胞内众多分子中分辨出不同荧光标记的结构。目前荧光技术正在稳步的发展，荧光标记物的数量和质量都在不断地提高。世界上最大的荧光物生产厂商 Molecular Probes 正在积极开发各种荧光标记物，其产品广泛应用于细胞生物学、神经科学、分子生物学及生物化学等领域，甚至在常规的蛋白质染色领域，目前开发的荧光染料正在取代传统的蛋白质染色金标准——银染和考马斯亮蓝染色法。

荧光、化学发光及生物发光技术具有安全、快速的检测特性，其发展呈现出检测微量化、高通量检测及操作简单化的趋势。表 8-2 列出了荧光与化学发光及生物发光技术的比较。

表 8-2　荧光与化学发光及生物发光技术的比较

项　　目	荧　　光	化学与生物发光
光强度	高	相对较低
背景	可变(取决于激发光波长与样品的组成)	低
光对样品的破坏	可变(取决于激发光波长与强度)	无
同时多点检测	多种选择	有限选择
成像	常规	困难

第二节　核酸的标记

20 世纪 50 年代开始的分子生物学研究与发展，直接来源于核酸结构研究的突破。DNA 双螺旋结构的发现是直接的导火索。随后的研究已经充分揭示了 DNA 与 RNA 分子在生命活动中的核心作用。通过对 DNA、RNA 分子进行标记，研究观察其在体内、体外各种不同活动过程中的变化规律，成为生物学领域中的一项基础工作。对 DNA 或 RNA 的标记方法很多，主要采用的是同位素标记方法；但近年来，标记方法开始逐渐发生了变化，各种非同位素标记方法日益成为重要的标记方法。

一、放射性同位素标记探针

探针是一段有特定序列的、有一定长度（15 个至近千个核苷酸）的核酸片段，也可以是完整基因，亦可以为其一部分；可以是天然序列，亦可以是人工合成序列。放射性同位素标记是核酸标记的常用方法。标记的核酸探针通过分子杂交来检查标本中的互补序列。目前最常用的同位素是 α-^{32}P dNTP、^{3}H-dNTP 及^{35}S-dNTP，标记方法多采用切口平移法、末端标记法、随机引物延伸法和逆转录标记法等。

（一）切口平移法

直到 20 世纪 70 年代初期，放射性同位素仅能通过代谢标记法掺入核酸。此法将放射性标记的前体（通常以 $H_3^{32}PO_4$ 的形式）引入正在合成靶 DNA 的细胞中，所需的放射性同位素用量之大，以现代更为保守的观念看来是难以置信的；放射性标记核酸的纯化费时费力，且只能用于少数 DNA（如病毒基因组）。20 世纪 70 年代中期发展了切口平移法，可进行体外高比活性 DNA 的放射性标记。从此，核酸的放射性标记研究获得了突破性的进展。

1. 基本原理

切口平移需要两种不同的酶活性参与反应。DNA 酶Ⅰ（DNase Ⅰ）可在双链靶 DNA 两条链上的随机位点切割（切口）磷酸二酯键，大肠杆菌 DNA 聚合酶Ⅰ则将脱氧核苷酸加到 DNase Ⅰ切割产生的 3′-羟基端。

在适当的浓度的 DNase Ⅰ作用下在一双链 DNA 上制造一些切口，双链 DNA 的单链切口在 DNA 聚合酶Ⅰ的 5′→3′外切活性作用下，从切口的 5′端逐个水解核苷酸时，酶的聚合活性则利用切口的 3′-游离羟基逐个加上相应的单核苷酸，使得切口向下游移动，这种切口移动的现象称为切口平移（见图 8-1）。如果反应中使用的单核苷酸底物是用同位素标记过的，则可合成高比活性的放射性标记的 DNA 分子探针。此反应所需的放射性前体为代谢标记法的 1/100 以下，而产生的双链探针比活性高于代谢标记法 200 倍。这样的探针可用于各种研究目的，包括基因组和 cDNA 文库的筛选、Southern 杂交和 Northern 杂交。

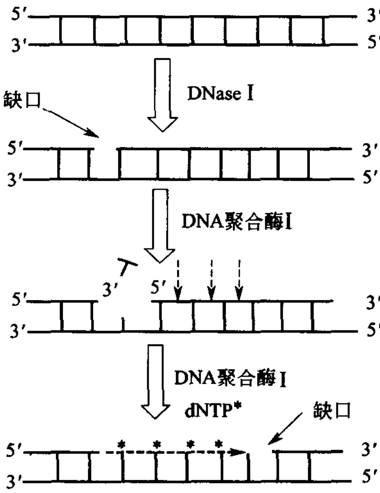


图 8-1 切口平移法制备 DNA 分子杂交探针
（＊表示同位素标记）

切口平移独领风骚多年之后，现已基本上被随机寡核苷酸引物法所取代，后者仅需一种酶参与反应，因而成为放射性标记双链 DNA 的首选方法。然而，随着切口平移试剂盒的诞生，标记过程也变为更简单有效。

2. 条件优化

切口平移反应中放射性标记物掺入的量取决于 DNase Ⅰ切割产生的模板 DNA 中 3′-羟基端的数量。切口过多会产生过多的起始位点，可导致放射性标记物的最大掺入，但却产生过短的 DNA 片段而不能用作杂交探针。另一方面，切口过少则使切口平移可用的起始位点数量受限，而产生低比活性的产物。制备过程中每批 DNase Ⅰ的活性可有所不同，大肠杆菌 DNA 聚合酶Ⅰ中所污染的 DNase Ⅰ的量亦相差悬殊，因此需要测定每批 DNase Ⅰ，摸索出酶的适当浓度，以产生所需比活性及长度的探针。

在合适的 DNase Ⅰ浓度下，反应中可得到约 40%的 [α-^{32}P] dNTP 掺入量。在此条件下，标准的切口平移反应可产生比活性高于 10^8cpm/μg 的放射性标记探针，标记 DNA 链的长度为约 400～750nt。因此，为了获得高比活性及长度适当的探针，可通过固定 DNA 聚合酶Ⅰ的用量而改变 DNase Ⅰ的用量，以确定每种酶的最佳用量，使反应得到优化。

dNTP 的用量也需要考虑。虽然大肠杆菌 DNA 聚合酶Ⅰ在高浓度底物存在时的催化效率最高，但为减少同位素用量，切口平移反应中通常含有最低浓度的放射性标记 dNTP (0.5～5μmol/L)，而含较高浓度的未标记的 dNTP（1mmol/L)。反应中的放射性标记的 dNTP 的比活性基本决定了终产物的比活性。这样，用同源未标记的 dNTP 来稀释放射性标记的 dNTP 可制备不同比活性的标记 DNA。如果需要制备高比活性（$>5\times10^8$dpm/μg）的探针（用于重组 DNA 文库的筛选或用于复杂的哺乳动物基因组 DNA 的 Southern 杂交检测单拷贝序列），切口平移反应中应含有所有 4 种放射性标记的 dNTP（比活性>800Ci/mmol)，且不含未标记的 dNTP。然而，对于大多数研究，使用一种 α-^{32}P（800Ci/mmol）标记的 dNTP 及 3 种未标记的 dNTP 即可满足要求，或用适量未标记的 dNTP 稀释每种

[α-^{32}P] dNTP。

3. 基本过程

此法也适用于探针的非放射性标记。通常的标记方法如下。

^{32}P 标记 DNA 探针（切口平移法）：反应体积为 25μl，内含 0.3μg DNA 片段，4μl 0.2μmol/L dNTP，1.1×10^6Bq [α-^{32}P] dATP，1μl 2 万倍稀释的 DNase Ⅰ和 2μl DNA 聚合酶Ⅰ，6μl 缓冲液 [为 50mmol/L Tris・HCl（pH7.2）、10mmol/L $MgSO_4$、1mmol/L 二硫苏糖醇（DTT）和 50μg/ml BSA]，反应在 14℃进行 3h。标记 DNA 经 Sephadex G-50 柱层析回收。

（二）末端标记法

1. T4 多核苷酸激酶催化的 5′端标记

催化核酸 5′端标记的首选酶是 T4 多核苷酸激酶（T4 polynucleotide kinase），该酶来源于 T4 噬菌体感染的大肠杆菌细胞。1985 年 C. Midgley 等成功地将编码该酶的基因克隆到大肠杆菌，获得了高效的表达。T4 多核苷酸激酶可将 [γ-^{32}P] ATP 的 γ-磷酸残基转移至去磷酸化的单链、双链 DNA 及 RNA 的 5′羟基端。当反应有效进行时，反应中 40%～50%的 5′突出端被放射性标记。但产生的探针不如其他放射性标记方法获得的比活性高，因为每个 DNA 分子中仅引入了一个放射性原子。尽管如此，用此方法合成的探针仍适合于多种研究目的。

除了直接将 ATP 的 γ-磷酸直接转移到去磷酸化的 DNA 的 5′-OH 末端外（这种反应被称为正向反应），T4 多核苷酸激酶还能够进行另外一种磷酸转移反应，称为交换反应。在交换反应中，过量的 ATP 可使 T4 多核苷酸激酶将磷酸化 DNA 的 5′端磷酸转移给 ADP，然后 DNA 从 γ-^{32}P-ATP 中获得放射性同位素标记的 γ-磷酸而重新磷酸化，如图 8-2 所示。正向反应一般比交换反应常用、有效。

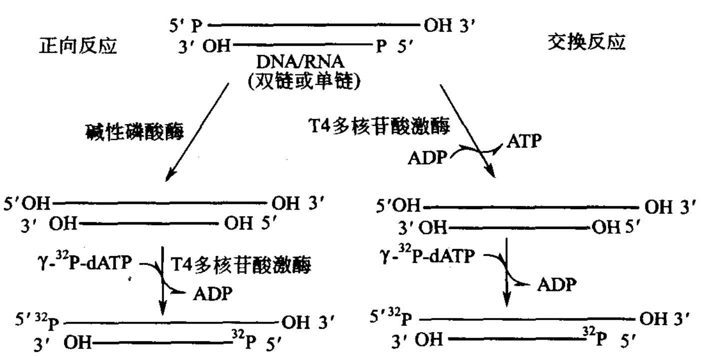


图 8-2 T4 多核苷酸激酶的正向反应和交换反应

在正向反应中单链或双链 DNA 首先用磷酸酶处理，以除去 5′-磷酸残基。然后，在 T4 多核苷酸激酶催化下利用 γ-^{32}P-ATP，重新使 5′端磷酸化，从而带上放射性标记。放射性标记的效率取决于 DNA 纯度（因为未纯化的 DNA 溶液含有激酶抑制剂）和 DNA 5′端的序列。由于未知的原因，5′端为胞嘧啶（C）的寡核苷酸片段的标记效率为以 A 或 T 开始的寡核苷酸片段低 1/4 以下，为以 G 开始的寡核苷酸片段的 1/6 以下。

标记方法如下。

将下列液体依次加入 Eppendorf 管中：

去磷酸化的 DNA	10～50pmol
10×T4 多核苷酸激酶缓冲液	5μl
10mCi/ml [γ-^{32}P]-ATP	
(比活性为 3000～7000Ci/mmol)	50pmol
T4 多核苷酸激酶	10 单位
水	加至 50μl

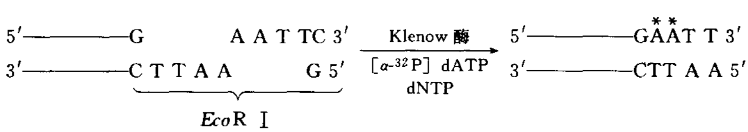
37℃温育反应 1h。

加入 2μl 0.5 mol/L EDTA (pH 8.0) 以终止反应。

用 Sephadex G-50 离心柱色谱分离放射性标记的 DNA。

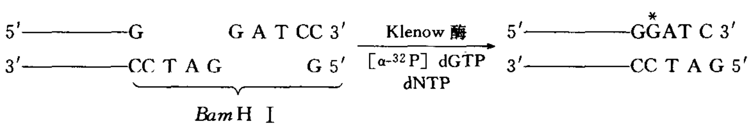
2. 大肠杆菌 DNA 聚合酶 I 的 Klenow 片段催化的 3′端标记

双链 DNA 的 3′端线性双链 DNA 最简单的标记方法是利用大肠杆菌 DNA 聚合酶 I 的 Klenow 片段催化一个或多个 [α-^{32}P] dNTP 掺入 3′凹端。对于平末端 DNA 则可用适当的限制性核酸内切酶消化，以产生适宜的模板片段，然后用 Klenow 酶催化 dNTP 补平 3′凹端。标记反应十分迅速，且反应能在限制酶消化液中进行，无需更换缓冲液。由于 Klenow 酶的 3′→5′外切酶活性较弱，因而对于平端或 3′突出端则几乎不能标记。根据需要，可选择标记线性双链 DNA 分子 3′的一端或两端。对平端的标记则需利用限制酶的识别与切割反应，以产生合适的 3′凹端。多数情况下，标记反应中最好含有未标记的 dNTP，并根据需要选择一种 [α-^{32}P] dNTP，因而允许在 3′凹端的任何位置进行标记。当然，[α-^{32}P] dNTP 的选择取决于 DNA 末端 5′突出端的序列以及实验目的。如用 *Eco*R I 切出的 DNA 末端可用 [α-^{32}P] dATP 标记。



* 代表放射性标记

如此产生的含两个相同放射性标记核苷酸的 3′末端片段虽然不适用于 Maxam-Gilbert 测序（因为在测序放射性自显影图上将获得双联带型），但这样标记的带放射性的 DNA 可用作分子质量标准或探针。同样的道理，由 *Bam*H I 切割 DNA 产生的末端可用 [α-^{32}P] dGTP 标记。



* 代表放射性标记

在上面的反应中，只有一个放射性标记的核苷酸掺入每一端。反应中加入未标记的 dNTP 可避免 Klenow 酶的 3′→5′外切酶活性切除模板 3′端核苷酸，并确保所有放射性标记的 DNA 产物长度相等。

有时，可通过选择适当的 [α-^{32}P] dNTP，实现只标记双链 DNA 分子的一端。例如，用 *Eco*R I 切开 DNA 的一端，而另一端用 *Bam*H I 酶切，则可在反应中加入 [α-^{32}P] dATP（标记 *Eco*R I 位点）或用 [α-^{32}P] dGTP（标记 *Bam*H I 位点）进行选择性标记。特殊情况下，如待标记 DNA 片段的末端序列是已知的，并有可利用的特殊的酶切位点，则可优先标

记 DNA 的某一端。例如，限制酶 *Hinf* Ⅰ 识别序列 5′-GANTC-3′，并在 G 和 A 之间切割产生 3′凹端。如果 DNA 片段 5′端的 *Hinf* Ⅰ 识别序列为 5′-GACTC-3′，而 3′端序列为 5′-GAGTC-3′，则可在反应中加入未标记的 dATP 和 [α-^{32}P] dGTP，选择性地在 5′端标记 DNA 片段，或加入未标记的 dATP 和 [α-^{32}P] dCTP，在 3′端选择性地标记 DNA 片段。限制性内切核酸酶如 *Dde*、*Fnu*4H、*Bsu*36 Ⅰ 和 *Eco*O109 Ⅰ 均可用于多种类似的标记策略。

上述两端均标记的 DNA 片段可用作 Southern 印迹中的分子量标准、凝胶阻滞试验的探针、凝胶上小量 DNA 的示踪剂、用于筛选菌落或噬菌斑以及用于重组质粒、质粒或 λ 噬菌体的 Southern 分析的低比活性探针。只标记一端的片段可用于供 Maxam-Gilbert 法测序用的底物、S Ⅰ 核酸酶或绿豆核酸酶进行 RNA 作图使用的探针、引物延伸反应中的引物等。

由 Klenow 酶催化的末端标记反应不限于放射性标记 dNTP，也适用于非放射性标记，如生物素、荧光素和地高辛修饰的脱氧核苷酸亦可使用。修饰的核苷酸掺入的 K_m 值通常低于标准 dNTP 的 K_m 值，这取决于标记物的性质。

3. 用 T4 噬菌体 DNA 聚合酶标记双链 DNA 的 3′端

T4 噬菌体 DNA 聚合酶比 Klenow 片段的 3′→5′外切核酸酶活性强得多，因此用 T4 DNA 聚合酶标记平端分子有其独到之处，尤其适合平端或 3′突出末端的标记。由于 T4 DNA 聚合酶的 3′→5′核酸外切活性对于单链底物的作用比双链底物更强，且比 Klenow 酶强约 200 倍。能迅速地消化 3′突出端，但对 DNA 底物双链部分的 3′核苷酸的水解速度较慢。在缺乏 dNTP 的条件下，此酶将双链 DNA 分子降解为约半长的单链。而在高浓度 dNTP 的存在下，核酸外切活性产生的凹进的 3′羟基端可作为引物，在 5′→3′聚合活性作用下，催化模板指导的单核苷酸添加。假如 4 种 dNTP 有一种被放射性标记，所产生的平端双链分子在靠近 3′末端处将被标记，如图 8-3 所示。限制性内切核酸酶如 *Pst* Ⅰ、*Pvu* Ⅱ 及 *Sac* Ⅰ 能不对称地酶切 DNA，产生带 3′突出端的片段，均可以用这一方法标记。

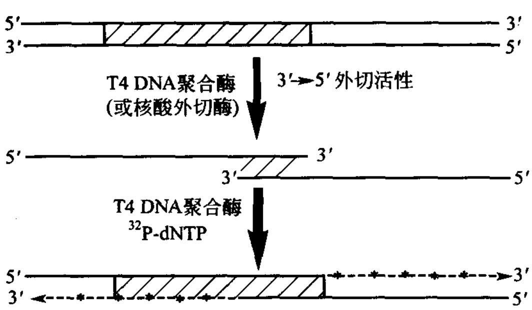


图 8-3 T4 DNA 聚合酶标记双链 DNA 的 3′端

对于外切核酸酶和 T4 噬菌体 DNA 聚合酶催化的补平反应的动力学的细致研究表明，放射性标记物系非随机地掺入 DNA。因此，典型的放射性标记片段含有大量新合成 DNA 尾部，而非整个 DNA 模板被末端标记。显然，这种标记反应模式与外切核酸酶的特性有关，即一旦反应开始，酶将沿 3′→5′方向酶解相当长度的 DNA，随后补平切去的 DNA。

应当注意的是，取代合成法不能够均匀标记整条 DNA 链，也难以标记片段中部的核苷酸，这一点与切口平移法完全不同。由于 T4 DNA 聚合酶的 3′外切酶活性并无序列特异性，在它的作用下，所有 DNA 片段长度减少的速度都基本相同；同时，由于 3′核酸外切酶的活性降解单链 DNA 的速度比降解双链 DNA 的速度快得多，因此当降解到中点时，一条限制片段将会分离成两个半长的单链，随后迅速地被完全降解。所以，及时终止外切反应对于成功的末端标记是十分重要的。

4. 末端脱氧核苷酸转移酶催化的 DNA3′端标记

末端脱氧核苷酸转移酶（terminal deoxynucleotidyl transferase），简称末端转移酶（terminal transferase），分子量为 60000，来源于前淋巴细胞和分化早期的类淋巴样细胞。

在二价阳离子存在下，该酶能够催化5′脱氧核苷三磷酸进行5′→3′方向的聚合作用，逐个地将脱氧核苷酸分子加到线性DNA分子的3′-OH末端。与DNA聚合酶不同，在反应中末端转移酶不需要模板的存在就可以催化DNA分子发生聚合作用。而且，它还是一种非特异性的酶，4种dNTPs中的任何一种都可以作为它的前体物，但对不同的核苷酸所需的阳离子略有不同。若加入的核苷酸为嘧啶核苷酸，则Co^{2+}为首选阳离子；若加入的核苷酸为嘌呤核苷酸，则Mg^{2+}为首选阳离子。

Pst Ⅰ、*Sac* Ⅰ或其他限制酶消化产生的DNA 3′突出端可通过牛胸腺末端转移酶催化[α-^{32}P]双脱氧ATP或[α-^{32}P] 3′脱氧腺苷三磷酸转移而被标记。由于上述两分子均不含3′羟基，因而其3′端不能连接其他核苷酸，而使合成终止。多数情况下，由于[α-^{32}P]双脱氧ATP比[α-^{32}P] 3′脱氧腺苷三磷酸的掺入更有效，因而成为3′端标记的首选底物。但是，3′端含3′脱氧腺苷残基的DNA分子有一个优点，即能抵抗污染的外切核酸酶的消化，从而防止3′端标记物的降解。

此法同样适用非同位素标记，如将底物改为bio-11-dUTP，也可以在3′端标记上一个生物素。

（三）随机引物法标记DNA片段

利用寡核苷酸随机引物，在DNA聚合酶催化下可沿单链模板起始DNA的合成。由于采用的是随机引物，因此可在模板的多位点上以同样的频率启动DNA的合成。用一种[α-^{32}P]标记的dNTP及其他三种非标记dNTP作为前体合成标记的DNA，可产生比活性为$5\times10^{8}\sim5\times10^{9}$ dpm/μg的探针。

与切口平移法相比，随机引物法更为简单，因为它不需要两种酶（DNase Ⅰ和5′→3′外切核酸酶）参与反应，而且随机引物反应产生的放射性标记产物长度更为均一，在杂交反应中重复性更好。切口平移法无法精确地控制放射性标记产物的平均长度，而在随机引物反应中，探针长度与引物的浓度呈负相关，因而探针DNA的平均长度可受研究者的控制。实验充分证实了随机引物法的优越性，它作为双链DNA探针的标准标记方法已完全取代了切口平移法。

（1）酶的选择　随机引物反应中可使用多种DNA聚合酶，但以Klenow酶为首选。因为Klenow片段缺乏5′→3′外切酶活性，这样放射性标记产物全部通过引物延伸而非切口平移合成，并且不会被外切核酸酶所降解。反应在pH6.6的条件下进行，此时该酶的3′→5′外切酶活性很低。此外，亦可应用Sequenase或Klenow酶的突变产物，此两种酶均缺乏3′→5′外切酶活性。

（2）放射性标记物　随机引物反应一般以一种放射性标记的[α-^{32}P] dNTP和三种未标记的dNTP作为前体。在适当反应条件下，40%～80%的[α-^{32}P] dNTP可掺入DNA，放射性标记产物的比活性范围在$1\times10^{9}\sim4\times10^{9}$ dpm/μg，它取决于反应中模板DNA含量。虽然用两种[α-^{32}P] dNTP作为前体可使标记产物的比活性提高，但由于放射性化学衰变可导致探针的迅速降解，因此标记后必须立即使用。当以低比活性的[α-^{32}P] dNTP作为前体时（如比活性为800Ci/mmol），DNA的产量可提高4倍，但比活性则降至<10^{9} dpm/μg，这一水平足以满足大多数实验的要求。在辐射分解造成探针降解之前，尚可冷冻保存数天。

（3）引物　随机引物的长度至关重要，短于6个碱基的引物是不可取的，而长于7个碱基的引物则增加了自身复性及自身引导的概率。因此，适用于随机引物法的理想寡核苷酸引物集群的长度应当为6个或7个核苷酸，集群中所有可能的序列应当以相同的频率出现。

随机引物反应合成的探针的平均长度与引物的浓度成反比，放射性标记产物的长度 $L=k/\sqrt{\ln P_c}$，P_c 表示引物的浓度。

例如，引物反应采用 60ng 和 125ng 六聚体或七聚体的随机寡核苷酸，产生的放射性标记产物长度为约 400～600bp，可用碱性琼脂糖凝胶或变性聚丙烯酰胺凝胶电泳测定探针的长度。引物浓度过高将使产量降低；较低则导致放射性标记产物集群大小不均一（0.4～4kb），而引起杂交动力学异常。

（4）模板 DNA　随机引物法可用于标记长度达 1kb 的线性双链 DNA 模板，较短的 DNA 模板产生的探针比活性低，在严谨条件下杂交效果不佳。闭环的双链 DNA 是低效模板，在进行随机引物反应之前应当用限制性核酸内切酶消化使之线性化。条件允许的情况下，应尽可能用纯化的 DNA 片段代替完整质粒作模板，这样的放射性标记 DNA 在用作杂交探针时可使本底水平显著降低。

（四）聚合酶链反应标记高活性 DNA 探针

聚合酶链反应（PCR）可用于制备高比活性的 DNA 探针。该法的优越性如下：①可扩增并标记目的 DNA 的限定片段，而不依赖于限制性酶切位点的定位与类型；②不需要分离 DNA 片段或将片段亚克隆入含噬菌体启动子的载体；③仅需要小量的模板 DNA（2～10 ng）；④放射性标记扩增的 DNA 的比活性可达 10^9cpm/μg。具体方法有 4 种，分述如下。

1. 应用常规 PCR 法获得双链 DNA 探针

PCR 反应中含有同样浓度的上、下游引物，三种未标记的 dNTP 浓度均超过各自的 K_m 值，一种［α-^{32}P］标记的 dNTP 浓度等于或略高于其 K_m 值（2～3μmol/L）。

2. 通过 PCR 合成以 DNA 的一条链占优势的放射性标记探针（不对称探针）

此反应中，一条引物的浓度为另一条引物的 20～200 倍。在 PCR 最初的几个循环中，双链 DNA 以常规指数方式扩增。但当另一条引物的浓度受限时，反应则以算术级数积聚单链 DNA，结果合成的 DNA 一条链的浓度大于另一条链 3～5 倍。

3. 通过热循环反应合成完全由 DNA 一条链构成的放射性标记探针

此反应中含有双链 DNA 模板，但仅含有一条引物。通过 40 个循环反应，双链 DNA 模板（20ng）可产生约 200ng 的单链探针。通过引物结合位点下游的一个限制性位点，切割模板 DNA，可确定探针的长度。

4. 利用 GC 切点的扩增反应

限制性核酸内切酶 *Cvi* JⅠ是以 ATP 为辅助因子，切割识别序列 GC（YGCR 除外，Y 为嘧啶，R 为嘌呤）的限制酶。由于二核苷酸 GC 在大多数 DNA 中频繁出现，因而酶切后产生分子量较小的平端 DNA 片段，长度为 20～60 个核苷酸。这些片段可作为 PCR 反应中的序列特异性引物，取靶 DNA 的一部分（约 1/10）作为模板，产生分子量不均一的双链 DNA 分子群。大小范围从最小 60bp 到最大超过靶 DNA 分子。这些大 DNA 分子被认为是自嵌合模板与引物产生的复杂而混乱的产物。这一方法的独特之处在于它不需要知道靶序列的序列信息。

二、非放射性同位素标记

自从 1981 年合成了生物素化 dUTP 和 NTP 类似物后，非同位素标记逐渐开始得到应用。目前，用生物素标记核酸探针检测和识别特定的核酸序列，已广泛应用于临床检验和科学研究。例如，生物素标记的 DNA 探针成功地用于 Southern 印迹杂交。随后，也发展了用化学

法直接把辣根过氧化物酶接到DNA上，运用发光显影检测的方法。用地高辛（digoxigenin）标记DNA，经地高辛抗体连接碱性磷酸酶，催化发光底物金刚烷脱磷并发光，其检测灵敏度与^{32}P标记探针相同，而且只需10～30min，即能快速检测，十分简便；而用传统的^{32}P标记的DNA探针杂交膜方法，需要－70℃曝光4～7d，才能在X射线胶片上显影。这显示了利用非同位素标记的化学或生物发光检测的良好应用前景与实用价值。非同位素标记与检测分为直接的和间接的方法。

（一）间接非同位素标记

前述放射性标记中，放射性原子只是简单地掺入探针的天然结构而取代非放射性同系物。在非同位素标记方法中，则是通过正常情况下核酸中不存在的化学基团在酶学、光化学或合成反应过程中与探针结合、反应，从而进行检测。如果报告酶并非直接结合于探针，而是通过桥联反应结合（如链亲和素-生物素），则这类非放射性检测被称作间接系统。

当探针与靶核酸杂交后，探针上的修饰基团可通过适当的指示系统（或报告酶）而被检测。20世纪80年代发展起来的第一个非放射性同位素检测方法是在用二硝基酚、溴脱氧尿苷或生物素标记核酸探针的基础上建立的。其中生物素具有重要的作用，杂交后生物素化探针与结合于报告酶（通常是碱性磷酸酶）上的链亲和素相互作用而被检测。

1. 生物素化核酸标记

生物素（biotin）是一种常见的维生素，它能与抗生物素蛋白（avidin）或链亲和素（streptavidin）特异性结合，而被用于各种检测。在间接检测系统中，利用生物素化的核酸（如bio-11-dUTP或bio-7-dUTP）作为标记试剂（中间数字表示连接臂的碳链长度），生物素部分通过一条4～16碳原子的间隔基与核酸结合，这样就减少了生物素化核酸探针在合成及检测中形成空间位阻的可能性。大多数情况下，生物素化的探针作用良好，不会发生问题。然而，由于生物素普遍存在于哺乳动物组织中，生物素化探针倾向于牢固地附着于某些类型的尼龙膜，因此在原位杂交、Northern杂交和Southern杂交中本底较高。

2. 地高辛标记

异羟基洋地黄毒苷，简称地高辛（DIG）。与生物素相类似，DIG也能通过化学方法偶联于接头和核苷酸，亦可通过任何一种标准的酶学方法，将DIG标记的核苷酸掺入核酸探针。一般来说，DIG标记的探针与各种类型的膜均具有低亲和性，因此这些探针比生物素标记探针的本底低得多。DIG标记的杂交体可用于报告酶偶联的抗DIG Fab片段检测（通常为碱性磷酸酶，较少的情况下采用辣根过氧化物酶）。

3. 荧光素标记

荧光素（fluorescein）标记的探针（荧光素-12-UTP或-dUTP）用于与碱性磷酸酶连接的抗荧光素抗体检测。用荧光素、生物素或DIG进行核酸的标记同样容易，而且以荧光素标记为基础的间接系统的敏感性至少与生物素和DIG系统相同。三种标记系统均能检测在Southern印迹和Northern印迹中亚皮克水平的靶核酸。尽管如此，由于某些不清楚的原因，荧光素标记的间接系统的应用不及生物素和DIG系统。目前3种系统均有商品出售，但DIG标记探针的非同位素检测系统只有Boehringer Mannheim公司一家销售。

（二）用生物素、地高辛或荧光素标记探针的方法

1. 酶学方法

利用生物素化、地高辛化或荧光素化的核苷酸标记DNA，可借助下述任一种标准的酶学技术：随机寡核苷酸引物法、PCR、切口平移或末端转移酶加尾法。通过SP6、T3或T7

聚合酶催化的体外转录反应可合成标记 RNA。这些反应中的标记物均为修饰的尿嘧啶核糖核苷三磷酸或尿嘧啶脱氧核糖核苷三磷酸，在标记反应中被用来取代 UTP 或 dTTP。对于大多数酶学标记反应，首选底物为 11 或 12 位上携带加合物的尿嘧啶三磷酸（如生物素-11-UTP 或 DIG-11-dUTP 或荧光素-12-dUTP）。而生物素-和 DIG-16-dUTP 是用于末端转移酶标记 3′端的偏爱底物。各种方法标记 DIG 的比较列于表 8-3 中。

表 8-3　非放射性 DIG 标记方法的比较

方　法	优　　点	模板的需要量	探针的检测灵敏度	应　　用
PCR 标记法	模板用量少，纯度不高的模板也可以进行标记；较其他标记方法而言，需要优化的条件少；标记探针的产量高；用于短探针（＜100bp）的标记，产生的探针非常敏感	质粒 DNA 10～100pg（使用 10pg 较理想）。基因组 DNA 1～50ng（使用 10ng 较理想）	0.10～0.03pg DNA	基因组 Southern 印迹 Northern 印迹 文库筛选、斑点/狭缝印迹、原位杂交
随机引物标记法	产生非常敏感的探针；反应体积能够十倍规模地扩大	300ng DNA（产生的探针可检测单拷贝基因）	0.10～0.03pg DNA	基因组 Southern 印迹 文库筛选 斑点/狭缝印迹 Northern 印迹
切口平移法	探针长度可控制（对于原位杂交很重要）	1μg DNA	敏感性取决于靶基因系统	原位杂交
化学标记法	技术简单，能够标记任意长度的探针，反应体积可扩大；可以标记线性和超螺旋质粒 DNA、总 RNA 和mRNA	0.2～5.0μg DNA/RNA	0.10～0.03pg DNA	标记总 RNA、mRNA 或 cDNA；大量 DNA 或 RNA 的标记
RNA 标记法	产生的探针量最大；产生的探针完全不受载体序列的影响；RNA 探针是单链的；在 Northen 印迹中，RNA 探针比 DNA 探针更敏感；DIG-标记的 RNA 探针易被切割成小片段，适用于原位杂交	线性化的质粒 DNA 1μg。带启动子的 PCR 产物 100～200ng	0.10～0.03pg DNA。 0.10～0.03pg RNA	Northern 印迹 原位杂交 文库筛选 斑点/狭缝印迹
3′末端标记	仅需少量模板；反应体积可以成倍扩大	100pmol 寡核苷酸	10pg DNA	文库筛选 斑点/狭缝印迹 原位杂交
3′加尾标记	仅需少量模板；产生的探针比末端标记更敏感；反应体积可以成倍扩大	100pmol 寡核苷酸	1pg DNA	文库筛选 斑点/狭缝印迹 原位杂交

放射性同位素的标记效率可通过探针的最终比活性很容易计算。相比之下，非同位素标记反应的进程或计算其标记效率尚无简单快捷的方法。在生物素和 DIG 标记探针的情况下，据厂家介绍，探针中不超过 30%～35%的胸苷残基被标记的尿苷残基取代。较高水平的置换可引起靶序列检测敏感性的降低，这可能是由于杂交步骤中紧密排列的加合物之间的空间位阻所致。然而，由于缺乏有效的监测系统，非放射性探针的合成最好设立预反应，在标记反应的过程中以一定时间间隔取样，在斑点或狭缝印迹中用不同稀释度的靶 DNA 试验标记产物。在掌握标记产物已具有所需的敏感度后，再设计及进行大规模标记反应。

2. 光标记

生物素和 DIG 可在光化学反应中附着于核酸。这两种标记物均结合于硝基苯叠氮基团，

后者经紫外线或强可见光照射可转变为具高反应性的物质，能与 DNA 和 RNA 形成稳定的共价键。光化学标记效率远低于酶学标记：最好的情况下，150 个碱基中仅有一个碱基被修饰，产生的探针在哺乳动物 DNA 的 Southern 分析中不足以检测单拷贝序列。

3. 杂交后非放射性标记探针的检测

用 Southern 印迹和 Northern 印迹杂交后，检测用生物素、荧光素或 DIG 标记的探针需要两个步骤。首先将膜暴露于合成的缀合物，该缀合物由报道酶和一个由能高度特异性地紧密结合于标记部分的配基组成。对于生物素，大多数检测方法是利用生物素和与酶缀合的链亲和素之间的高亲和反应。也可选择抗生物素抗体。荧光素化或地高辛化探针一般通过与特异性的酶联抗体的相互反应而被检测。然后，再用比色、产荧光或化学发光的底物来检测报道酶。

在比色检测中，两种染料组合用于检测被标记的杂交体捕获的碱性磷酸酶。该酶催化自 BCIP（5-溴-4-氯-3-吲哚磷酸）除去磷酸基团产生的产物能氧化及二聚化形成二溴二氯靛蓝。在二聚化反应中产生的还原等价物，可将 NBT（氮蓝四唑）还原为不溶性的紫色染料，从而使标记探针与其靶位杂交的位点变成可见的。观察结果可通过肉眼分析或常规拍照。比色检测用于硝酸纤维膜和 PVDF 膜效果良好，但在尼龙膜及荷电的尼龙膜上则略有逊色。即使是在最佳情况下，比色法仍比其他非同位素检测系统（如化学发光法）的敏感性低两个数量级或更差。

碱性磷酸酶的荧光测定可利用 HNPP（2-羟-3-奈酸-2′-苯基苯胺磷酸）进行。去磷酸化后，HNPP 在膜上产生荧光沉淀，可被 290nm 的紫外线激发。利用与常规溴化乙锭染色凝胶相同的方法拍照或记录。利用 HNPP 的荧光分析比比色分析更敏感，而膜与底物温育 2～10h 则能检测哺乳动物基因组 DNA Southern 印迹中的单拷贝序列。荧光沉淀可通过乙醇洗涤很容易地去除，还可反复使用多次。

化学发光法利用含辣根过氧化物酶（HRP）或碱性磷酸酶的缀合物检测生物素、荧光素或 DIG 标记的 DNA，是最迅速、最灵敏的检测方法。HRP-鲁米诺检测系统由 Amersham 公司开发。HRP 在过氧化氢存在下能催化鲁米诺的氧化，产生高反应性的内过氧化物，在分解至基态时发射 425 nm 光。范围广泛的化合物包括苯并噻唑、酚、萘酚和芳香胺能显著增强反应中的发光量。这种高达 1000 倍的信号放大增强了 HRP-鲁米诺系统的灵敏度，达到能检测哺乳动物基因组 DNA Southern 印迹中的单拷贝基因的程度。反应的最初几分钟即能迅速地产生光信号，其后减弱，1～2h 后即不再能检测到。在理想的情况下，用蓝光敏感的浅灰色 X 射线胶片曝光 1h 或用冷 CCD 照相机可在 Southern 印迹中检测出约 1 pg 靶 DNA。于 80℃加热可从滤膜上剥离探针。

以二氧杂环丁烷为基础的化学发光底物是现有的检测碱性磷酸酶的最灵敏的化学发光底物。碱性磷酸酶切割底物的磷酸酯，使二氧杂环丁烷环的热稳定性显著降低，随后因 465 nm 的光辐射而分解。尼龙膜（两性电荷或正电荷）通过提供疏水域使信号明显增强，化学发光反应中产生的去磷酸化中间体在这些疏水表位螯合至前述疏水域。这使中间体得以稳定，并降低其非发光性分解。结果，激发的二氧杂环丁烷阴离子发射光的强度作为辉光开始，强度在几分钟内增加，以后可持续几小时。尼龙膜和阴离子之间的疏水相互作用可引起发射光从 477nm 至 466nm，约 10nm 的“蓝移”。

在多数实验情况下，化学发光物在尼龙膜上的延长动力学是有益的，可为多次曝光的影像捕获赢得时间。在理想的条件下，低至 1×10^{-19}mol 的 DNA 亦可检测出。这意味着当尼

龙膜在浅灰色蓝光敏感的 X 射线胶片上曝光约 5min 或用冷 CCD 相机即可检测到哺乳动物基因组 DNA Southern 印迹中单拷贝基因。在 65℃用 50%甲酰胺处理可除去探针，尼龙膜则可再用于几轮杂交。

（三）DIG 标记实例

1. DIG 的 PCR 标记法

PCR 标记法可以从非常少的模板产生大量的标记探针。PCR 扩增方法时，在底物中添加或掺入一定比例的 DIG-dUTP，扩增的产物片段即可被标记。模板的量从 10～100 pg 质粒或 1～50 ng 基因组 DNA 均可，PCR 标记反应中的其他成分见表 8-4。PCR 引物的特异性决定了扩增和标记的区域。这种方法特别适合于合成短的探针，用于基因组 Southern 印迹中检测单拷贝基因序列和 Northern 印迹中稀有 mRNA 的检测，另外也可用于文库筛选、斑点/狭缝印迹和原位杂交。PCR 标记反应体系见表 8-3。循环参数为：95℃预变性 2min；95℃变性 30s，60℃退火 30s，72℃延伸 40s，10 个循环；接着，95℃变性 30s，60℃退火 30s，72℃延伸 40～60s（视片段长短而定），20 个循环；最后 72℃延伸 7min。

表 8-4 PCR 标记反应体系组成

成　分	DIG 标记的对照管	未标记的对照管	终浓度
灭菌水	适量	适量	
10×PCR 缓冲液(含 $MgCl_2$)	5μl	5μl	
10×PCR DIG 混合物	5μl	—	200μmol/L dNTP
10×dNTP 储液	—	5μl	200μmol/L dNTP
上下游引物	适量	适量	每种引物 0.1～1μmol/L
酶混合物	0.75μl	0.75μl	2.6U
模板 DNA	适量	适量	10ng 基因组 DNA 或 10pg 质粒
总反应体积	50μl	50μl	

2. 随机引物标记法

该法采用随机六核苷酸引物，可标记任意长度的模板，产生出同源性高、敏感性好的杂交探针，能够检测出 0.1～0.03pg 的靶 DNA 分子，适合于基因组 Southern 印迹检测单拷贝基因和重组文库的筛选，也适合于斑点/狭缝印迹和 Northern 印迹。通常，由 2kb 模板产生的标记探针将分布在大约 300～1500bp 之间，大多数约为 800bp 左右。与 PCR 标记方法不同，随机引物标记法对模板的纯度要求较高。

下列反应步骤可标记 10ng～3μg 的纯化 DNA 模板。

①反应管中加入模板 DNA，加无菌水至 16μl；②沸水中加热样品，变性 10 min，然后迅速置于冰上；③在反应管中加入 4μl 混合物（随机引物、dNTP、DIG-dUTP、Klenow 酶、缓冲液），离心混匀，37℃温育 1～20h（如果模板量少，则温育时间可较长）；④加入 2μl 0.2mol/L EDTA（pH8.0）终止反应。

3. 体外转录标记 RNA 探针

选择具有启动子的载体克隆目的基因，则可获得 RNA 探针。如果在反应体系中加入 DIG-dNTP 则可标记转录生成的 RNA 探针。在标准反应体系中，1μg DNA 可产生 10～20μg RNA，大约每 25～30 个核苷酸可掺入 1 个 DIG 标记的 UTP。由于标记的 RNA 探针非常敏感，因此它更适合 Northern 印迹中稀有 mRNA 的检测，也适合于 Southern 印迹、文库筛选、斑点/狭缝印迹及原位杂交。但实验中要注意防止 RNA 酶的污染。体外转录标记 RNA 探针方法所需要的模板可通过克隆的质粒转录获得，也可通过 RT-PCR 方法制备。

分述如下。

（1）通过质粒转录方法

① 纯化的靶 DNA 插入到转录载体的多克隆位点，位于 RNA 聚合酶启动子的下游区。

② 在大肠杆菌细胞中扩增质粒，使用 CsCl 密度离心法纯化质粒，或利用质粒纯化试剂盒来提取质粒。

③ 用限制性内切酶切割质粒，使之线性化，切割后为 5′黏性末端突出。如果模板为 3′黏性末端突出或产生平末端则转录反应会产生“发夹”结构的产物，降低检测的敏感性，增加背景。

④ 酶切反应后，用酚-氯仿抽提纯化模板或用 PCR 产物纯化试剂盒来纯化模板，纯化后的模板溶在灭菌的 DEPC 水中或 10mmol/L Tris（pH8.0）中，浓度为 200～250μg/ml。

⑤ 使用 1μg（4～5μl）纯化的线性质粒作为模板进行标记反应。

（2）通过 RT-PCR 的方法准备 DNA 模板

① 用常规的方法提取总 RNA。

② 以 oligo（dT）为引物把总 RNA 进行逆转录反应。

③ 合成一对 PCR 引物，包含有与靶序列（即将被转录的序列）的互补部分和任意一种 RNA 启动子的互补序列（如 T7 启动子的互补序列）。

④ 在无菌离心管中依次加入下列成分，冰浴上操作。

成　分	体　积	终　浓　度
灭菌 DEPC 处理水	适量	
10×PCR buffer	5μl	1×，含 1.5mmol/L $MgCl_2$
dNTPs(每种 10mmol/L)	1μl	200μmol/L dNTP
上、下游引物	适量	300nmol/L
Taq 酶	0.75μl	2.6U
cDNA	2μl	
总反应体积	50μl	

⑤ 离心混匀后，按下列循环参数进行 PCR：94℃预变性 2min；94℃变性 45s，60℃退火 45s，72℃延伸 90s，30 个循环。反应结束后取 4μl PCR 产物（约 100～200ng），进行标记反应。

（四）直接非同位素标记

用荧光化合物直接标记核酸探针进行检测的方法称为直接检测系统。可直接标记使用的荧光染料有荧光素、得克萨斯红或罗丹明，镧系元素螯合剂，吖啶酸酯；也可直接共价结合碱性磷酸酶或辣根过氧化物酶用于检测。检测可采用比色法、化学发光法、生物发光法、时间分辨荧光测定术或能量转移/荧光猝灭等多种技术的任何一种。如利用特异性结合的抗体，则可将直接检测法转变成间接检测法。

用荧光化合物直接标记核酸为自动化 DNA 测序奠定了基础，使分子生物学发生了革命性的变化。但在其他方面，直接标记应用较为有限，其原因为：①用酶（如碱性磷酸酶）直接偶联的传统方法需广泛采用柱色谱、聚丙烯酰胺凝胶电泳与/或 HPLC 以获得纯化产物，操作过程复杂，成本价格昂贵，且产物不纯；②大分子的酶如碱性磷酸酶一经偶联，则可使探针复性的速率降低，在通常用于核酸复性的条件下可能不稳定，因此增加了操作难度；③本底高、灵敏度有限。

由于存在这些问题，直接检测技术已经失宠，而间接检测技术已经成为 Southern 杂交、Northern 杂交和原位杂交中靶核酸定位的主导的非放射性检测方法。然而，尽管存在这些问题，用碱性磷酸酶直接标记的探针现仍有商品生产，用于法医学和病理学分析。这些探针的灵敏性接近同位素标记的探针。

第三节　蛋白质的标记

蛋白质是重要的生物大分子，它与核酸一样是分子生物学研究中的主要对象。蛋白质的标记与核酸的标记目的是一样的，都是为了能够观察、检测相应的分子或作用过程，但两者的标记方法与原理明显不同：核酸标记在很多情况下利用了核酸分子的合成过程，对合成所需的原料（4 种 dNTP）进行标记，使之掺入核酸分子；蛋白质标记则不同，它只能利用蛋白质分子中氨基酸残基的特定基团，通过基团与标记物的反应而使蛋白质带上特定的标记。通过标记的蛋白质分子可以观察、检测与其相互作用的其他分子或特定的生物结构。蛋白质的标记主要采用非同位素的标记方法，包括荧光标记、化学及生物发光标记等。

一、荧光标记

荧光是指一个分子或原子吸收了给予的能量后，即刻引起发光；停止能量供给，发光亦瞬即停止。蛋白质标记中所用的荧光染料种类繁多，性能各异，标记方法也不尽相同。常根据荧光染料标记时所利用的结合的基团分为氨基反应染料、巯基反应染料、羟基反应染料及酰基反应染料等。它们涵盖了荧光素类、罗丹明类、二苯乙烯类、萘酰亚氨类、香豆素类等多种荧光染料。以下介绍各种不同荧光染料标记时常用的连接方法。

（一）利用氨基的标记方法

利用生物分子中氨基的标记反应是最常见的标记方法。标记的对象包括蛋白质、肽、配体、寡核苷酸等。氨基反应产生的生物分子连接物主要应用于免疫化学、荧光原位杂交、细胞示踪、受体标记、荧光类似物的细胞化学。在这些研究中，荧光染料与生物分子之间结合的化学键的稳定程度至关重要，它必须能经受较长时间、反复的观察研究，并能在经过剧烈的杂交与洗涤步骤后仍然保持稳固的连接。氨基反应染料种类较多，涵盖整个可见光的检测波长。

1. 制备合适的生物连接物

理想的生物连接物通常具有较高的荧光标记量，并同时保留了未标记生物分子的关键性能参数，如对受体的选择性结合能力、对酶的抑制或激活能力或掺入生物膜的能力等。染料与生物分子的结合通常是相当容易的，但是最大程度的结合常导致结合物沉淀或非特异性结合能力的增强。因此，标记时有必要降低标记的量，以保留其功能及特异性。对具体的标记反应，应通过实验确定最佳的标记量或参考试剂厂商提供的标准。标记过程一般并不复杂，也不需要特殊的设备。标记后务必将连接的产物与标记染料分开，因为游离染料的存在将严重影响后续检测过程及结果。通常，可通过灵活采用凝胶过滤或透析的方法分离标记物。

2. 氨基基团的标记反应

氨基反应性染料大多数是酰化试剂，它们能与氨基形成甲酰胺、磺酰胺、尿素或硫脲等基团。反应过程或动力学取决于酰化试剂与氨基的浓度和反应性。显然，含游离氨基的缓冲溶液如 Tris 等，会干扰标记过程，应当避免使用。

影响氨基标记反应的主要因素是其种类和碱性强弱。事实上，所有蛋白质含有赖氨酸残基以及大多数蛋白质含有 N 端游离氨基。脂肪族氨基如赖氨酸中的 ε-氨基，有一定碱性，能与大多数酰基试剂反应。然而，在低于 pH8 的条件下，以游离碱型存在的脂肪族氨基的浓度很低，这样在与异硫氰酸盐、琥珀酰亚胺酯等酰化试剂进行氨基酰化反应时，反应的过程或动力学严重依赖于溶液的 pH。对赖氨酸而言，pH8.5～9.5 是较为理想的反应条件。相比之下，N 端的 α-氨基时常需要在接近中性的环境中酰化。应当注意的是，尽管酰化反应可以在碱性条件下进行，但随着 pH 的升高，水溶液中的酰化试剂会逐渐降解。蛋白质与琥珀酰亚胺酯的反应通常在 pH8.5 的条件下进行，与异硫氰酸盐的连接则在 pH9.0 以上，这对于碱敏感的蛋白质是一个限制因素。

芳香族氨基在生物分子中并不常见，其碱性较弱，在 pH7 时未质子化。与芳香族氨基的反应需要高反应性试剂，如异硫氰酸盐、磺酰氯、酸性卤化物，但反应可在 pH4 以上的任何 pH 进行。酪氨酸残基可修饰形成一个 *O*-氨基酪氨酸芳香族氨基，它能在相对低的 pH 条件下与某些氨基反应性试剂反应。

在水溶液中，酰化试剂实际上不能与肽键中的氨基、谷氨酰胺和天冬酰胺中的氨基、精氨酸中的胍基、组氨酸中的咪唑基、核酸中的非碱性氨基等发生反应。

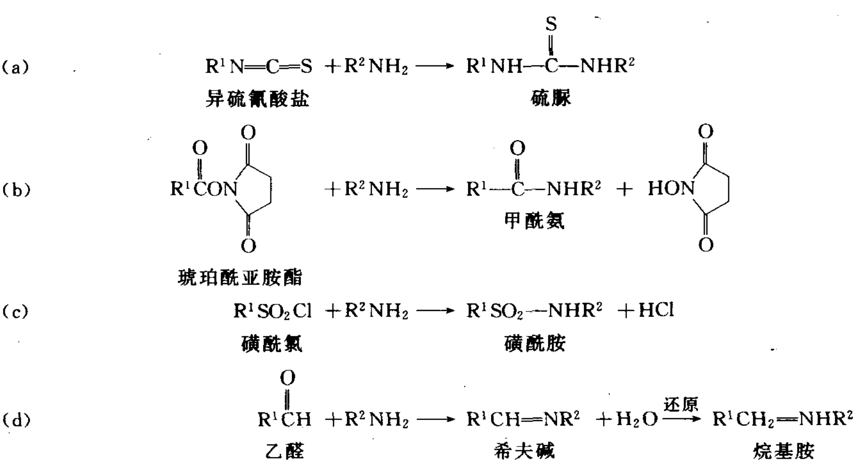


图 8-4　氨基与（a）异硫氰酸盐、（b）琥珀酰亚胺酯、（c）磺酰氯、（d）乙醛的连接反应

3. 具体实例

（1）异硫氰酸盐　异硫氰酸盐实际并不稳定，在储存过程中容易降解，因此现在已经较少直接应用。然而，有些易转化为异硫氰酸盐的酰基叠氮试剂在水及大多数溶液中相当稳定，可被用来替代异硫氰酸盐的标记。异硫氰酸盐与氨基反应后生成硫脲（图 8-4），该产物相当稳定，但有报道表明用异硫氰酸荧光物制备的交联抗体，在存放时会有部分降解。目前，异硫氰酸荧光素（fluorescein isothiocyanate，FITC）和四甲基罗丹明异硫氰酸盐（tetramethyl-rhodamine isothiocyanate，TRITC）仍然是交联抗体所使用的最常用的标记染料。

（2）琥珀酰亚胺酯　琥珀酰亚胺酯是标记氨基的极好染料，它与氨基形成的化学键与肽键一样稳定，在良好干燥条件下存放时极其稳定。它能与脂肪族氨基反应，而与芳香族氨基的反应性较低，也难与醇、酚、酪氨酸及组氨酸反应。琥珀酰亚胺酯在有机溶剂中能与硫醇类物质反应形成硫酯。尽管琥珀酰亚胺酯在交联反应中伴随着水解副反应，但在低于 pH9 的条件下，水解反应非常缓慢。当然，有些琥珀酰亚胺酯物质在水中溶解度低，难于用于标记反应。为了克服这不足，有公司开发了能转变成琥珀酰亚胺酯的衍生物，如磺化的试剂，其水溶性很强，以致不需要在有机溶剂中进行交联反应。磺化的琥珀酰亚胺酯类物质的纯化比较困难，这影响了其应用，目前这类荧光染料一般用于原位标记。

（3）磺酰卤化物　磺酰卤化物包括丹酰、罗丹明 B 及得克萨斯红衍生物等。这些物质在水溶液中，尤其是标记反应所需的较高的 pH 条件下很难溶解。例如，室温下得克萨斯红磺酰氯在 pH 8.5 的水溶液中 2～3min 即完全水解；但是，一旦交联，则生成的磺酰氨基极其稳定，以致在蛋白质完全水解时都保持连接状态。除了与氨基反应以外，磺酰氯也可与酚羟基（包括酪氨酸）、脂肪醇（包括多糖）、巯基（包括半胱氨酸）和咪唑基（包括组氨酸）反应，但这些反应在蛋白质及在水溶液中并不常见。与磺酰氯相比，磺酰氟的反应活性要低得多，这使得后者在某些酶的活性丝氨酸残基修饰反应中有特别的用途。应当注意，磺酰氯在 DMSO 溶液中不稳定，应避免使用。

（4）其他氨基标记试剂　醛类和酮类染料与氨基反应形成希夫碱。含醛基的标记试剂有荧光胺、邻苯二亚甲基二醛（OPA）、萘二羧醛（NDA）等。这些试剂对溶液中氨基的定量检测方面极其灵敏，可用于 HPTC 和毛细管电泳。

（二）利用巯基的标记方法

与氨基的标记反应不同，巯基标记反应主要通过标记肽、蛋白质、寡核苷酸等来探测生物学结构、功能及相互作用。由于可以高度选择性地标记巯基，故可通过激发态能量转换装置测定分子间或分子内的距离，以及通过巯基化引物来进行 DNA 测序，或含巯基蛋白的电泳检测等。

1. 巯基标记反应

蛋白质中巯基存在于半胱氨酸中，可以通过巯基试剂，如二硫苏糖醇（dithiothreitol, DTT）、二硫赤藓糖醇（1,4-dithioerythritol, DTE）、2-巯基乙醇（2-mercaptoethanol）选择性还原二硫键获得。然而，在标记反应前，这些巯基试剂必须除去。除去这些试剂的过程中（如常用的透析或凝胶过滤的方法）常导致巯基的氧化，重新形成二硫键。改进的方法是使用不含巯基的还原试剂，这样在标记巯基时就可不必除去还原试剂，避免巯基的氧化。有些试剂也可以将巯基引入蛋白分子或其他分子，这种方法在交联生物分子中特别重要。

巯基标记试剂主要是些烷化剂，包括碘乙酰胺类（iodoacetamides）、马来酰亚胺类（maleimides）、苯亚甲基卤化物（benzylic halides）及溴甲基酮类（bromomethylketones）。在室温或低于室温的条件下，及生理 pH 范围（6.5～8.0）内，它们与巯基的反应非常迅速，并产生稳定的硫酯类物质。

2. 碘乙酰胺类

碘乙酰胺类易与各种巯基反应，形成硫酯类物质（图 8-5）。碘乙酰胺类甚至比溴乙酰胺类更容易反应。然而，当蛋白质中没有半胱氨酸或半胱氨酸受到阻碍而无法接触时，它们可与甲硫氨酸反应，或可能与组氨酸、酪氨酸反应，但这些反应通常是在缺乏游离巯基时发生。碘乙酰胺类对光敏感，尤其在溶液中见光易分解，因此标记反应应该在避光条件下进

行。在反应溶液中加入半胱氨酸、谷胱甘肽、巯基琥珀酸能猝灭荧光探针，并形成高度水溶性的加合物，能通过透析或凝胶过滤很容易地移去。

3. 马来酰亚胺类

马来酰亚胺的化学名称是顺丁烯二酰亚胺，是一类极好的选择性巯基修饰、定量及分析的荧光染料。反应涉及巯基与双键交联形成硫酯（图 8-5）。它与甲硫氨酸、组氨酸及酪氨酸不发生反应，但其涵盖了碘乙酰胺类的应用范围。马来酰亚胺与氨基的反应所需要的 pH 条件要高于它与巯基反应所需的 pH。马来酰亚胺的水解产物无标记作用，但能与马来酰亚胺竞争巯基，这在 pH8 以上尤为突出。而且马来酰亚胺标记的加合物也能水解，并导致标记交联物的荧光特性显著改变，这些标记物似乎更适合用于巯基定量。

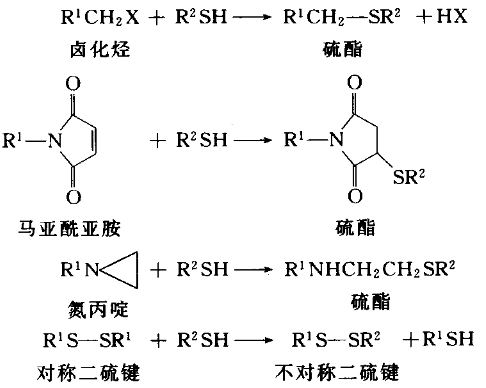


图 8-5 与巯基的反应

其他巯基标记试剂有烷化剂（烷基卤化物、氮丙啶、环氧化物等）、芳基化试剂（氟代硝基苯并呋喃、氧杂二唑磺酰胺等）。氮丙啶和环氧化物类通过将环打开形成硫酯类物。对称的二硫键化合物也可用于巯基标记反应，它通过可逆的反应与待标记的巯基交换而形成不对称的二硫键化合物（图 8-5）。这些标记的产物一般比上述的碘乙酰胺类和马来酰亚胺类的标记物更耐水解。

（三）利用醇羟基的标记方法

生物分子中的醇羟基是相当丰富的，如蛋白质中丝氨酸、苏氨酸及酪氨酸含有酚羟基，糖类中也含大量的羟基。但羟基并不十分活泼，它在水溶液中的反应性较低。因此，对蛋白质中丝氨酸、苏氨酸及酪氨酸中羟基的标记较为困难，除非这些羟基基团位于活性部位，具有较强的反应性。

有些磺酰氟化合物，包括甲苯磺酰氟（PMSF）和酰化剂，能与活性的丝氨酸残基反应。有报道用蒽酮酰腈（9-anthroylnitrile）标记肌球蛋白表面的某些丝、苏氨酸残基。对蛋白质 N 端的未酰化的丝氨酸残基、苏氨酸残基，可采用先将其氧化成醛，随后用肼类衍生物标记。对酪氨酸残基的标记有时是利用了磺酰氯和碘乙酰胺等其他反应性染料与蛋白质标记反应的副反应。例如，氯代硝基苯并呋喃与单胺氧化酶中活性酪氨酸残基反应。此外，单独对蛋白质中的酪氨酸残基标记可主要采用下述 2 种方法：①用四硝基甲烷对酚环的邻位进行硝基化，然后用连二亚硫酸钠还原邻硝基酪氨酸，形成邻氨基酪氨酸。后者在 pH≤7 的环境中能与大部分的氨基标记染料反应（琥珀酰亚胺酯除外）。这种间接的标记方法在用丹酰氯等标记肌动蛋白中已经得到应用。②用芳香重氮基盐与酪氨酸反应产生一种含氮染料加合物，后者同样可被连二亚硫酸钠还原成邻氨基酪氨酸，进而采用氨基标记方法。

二、利用酶、酶的底物及酶抑制剂的标记

目前，已有大量的产荧光的酶底物和产各种颜色的酶底物。利用这些底物可以检测出固定细胞或细胞提取物中的极低水平的酶活性，也能在酶联免疫检测（ELISA）中应用，同时还可以有效地检测到活细胞中的酶活性。底物种类很多，但对于不同的用途，对底物性能的要求也不一样。用于溶液检测或细胞提取液检测需要可溶性荧光底物，并希望其经酶作用

后，产物的荧光强度高、水溶性好，而且光谱特性与原底物显著不同，这样易于检测；用于活细胞中酶活性评价的底物则要求具有亲脂特性的底物或某些巯基反应性底物，亲脂性底物易于进入细胞并在脂溶性区域保持较高的浓度，巯基反应性底物在细胞膜内由谷胱甘肽-S-转移酶催化产生不能穿透膜的谷胱甘肽-荧光染料加合物；用于原位检测的底物要求其能产生不溶性的荧光产物，它能在组织化学染色、原位杂交以及 Western 斑点杂交、Southern 斑点杂交、Northern 斑点杂交等检测中将荧光染料固定于待检测的局部。

（一）糖苷酶及其底物

糖苷酶对相应的糖具有极高的选择性水解活性，例如 β-半乳糖苷酶能快速水解 β-D-吡喃（型）半乳糖苷，但不水解 α-D-吡喃（型）半乳糖苷或 β-D-吡喃（型）半乳糖苷的异构体。内源性糖苷酶的检测常用于鉴别微生物的株系、选择性标记哺乳动物细胞的细胞器等。此外，糖苷酶在报告基因检测方面特别重要，尤其是编码 β-半乳糖苷酶的 *LacZ* 基因。它由 IPTG［isopropyl β-thiogalactopyranoside，异丙基-β-巯基吡喃（型）半乳糖苷］诱导表达，其用于基因表达的发光底物通常是 X-gal［5-Bromo-4-chloro-3-indolyl β-galactopyranoside，5-溴-4-氯-3-吲哚-β-吡喃（型）半乳糖苷］，在酶作用部位产生深蓝色沉淀，这一反应在检测细胞或组织中 *LacZ* 活性等方面有重要用途。此外还有 5-溴-6-氯-3-吲哚、6-氯-3-吲哚、5-吲哚、*N*-甲基-3-吲哚以及吲哚取代的 β-D-吡喃（型）半乳糖苷等多种发光底物。

β-半乳糖苷酶也有多种荧光反应性底物，其中常用的是荧光素二-β-吡喃半乳糖苷（fluorescein di-β-galactopyranoside，FDG），非荧光性的 FDG 经 β-半乳糖苷酶水解后产生荧光素单-β-半乳糖苷（fluorescein monogalactoside，FMG），进而水解成强荧光物——荧光素（fluorescein）。这种荧光基础的底物检测极其灵敏，有报道表明使用 FDG 的酶联免疫（ELISA）其灵敏度比同位素标记的高 100～1000 倍。除了用于 ELISA 外，FDG 也可用于荧光显微镜下鉴定 *LacZ* 阳性细胞与用于流式细胞术，当然还可用于其他有关的检测中。

（二）磷酸酶类及其底物

磷酸与磷酸酯在细胞内被广泛用作酶的底物、第二信使、膜结构组分及储能物质。目前广泛应用的催化磷酸酯产生荧光或颜色物质的酶主要是简单的磷酸酶类，如碱性磷酸酶和酸性磷酸酶，两种酶都能水解磷酸单酯产生无机磷酸和醇，来源于小牛肠的碱性磷酸酶被广泛用于 ELISA、免疫组化技术、Western 斑点杂交、Southern 斑点杂交、Northern 斑点杂交分析。此外，磷酸酶也可作为酶标，以鉴定原始的干细胞、区分骨髓基质细胞的亚群以及研究肿瘤细胞的分化等。

在磷酸酶的产荧光底物中，二磷酸荧光素（fluoroscein diphosphate，FDP）是最敏感的产荧光磷酸酶底物之一。FDP 本身是无色的非荧光物，但被酶水解后则生成强荧光物。这使得 FDP 成为 ELISA 反应中碱性磷酸酶的极好底物。此外，FDP 在细胞黏附与迁移评价中用于检测内源性磷酸酶活性也极为敏感。Molecular Probes 公司研究了能在酶促反应作用部位产生荧光沉淀的底物，利用这种底物的检测被称为酶标荧光（enzyme-labeled fluorescence，ELF）技术。这种 ELF 磷酸酶底物本身具有微弱的蓝色荧光，经磷酸酶水解后，产生强的绿色荧光沉淀，其荧光强度比直接用荧光探针检测高 40 多倍，并且工作 pH 范围宽，既能检测酸性磷酸酶，又能检测碱性磷酸酶，适用于多种用途，包括流式细胞术和激光扫描共聚焦显微镜观测。

产颜色的磷酸酶底物常用的是吲哚磷酸类物质，如 5-溴-4 氯-3-吲哚磷酸（BCIP）。这种吲哚类物质在磷酸酶水解后，经氧化产生蓝色沉淀，因此也可用于原位检测。3-吲哚磷酸

的衍生物具有相似的成色反应，产生各种不同的颜色，如红紫色、粉红色、蓝色等，它们不溶于水、乙醇及二甲苯等，可以与苏木紫染料一起使用。

（三）其他酶与底物

糖苷酶与磷酸酶及其荧光底物是目前主要的酶活性检测系统，此外还有一些其他的酶检测系统，如氧化酶、脂肪酶、蛋白酶、肽酶等。另外，氯霉素乙酰基转移酶（chloramphenicol acetyltransferase，CAT）和荧光素酶（luciferase）也是较为重要的发光检测酶类。

CAT的活性非常灵敏，并且其活性与转录水平密切相关，因此成为研究动植物细胞基因表达研究的良好的报告基因。以前CAT的活性检测通常利用放射性同位素标记底物，如用 [^{14}C]-氯霉素或 [^{14}C]-乙酰CoA，经有机溶剂提取后闪烁记数或放射自显影检测。现在已有很多公司都生产了其荧光检测的方法。通过荧光氯霉素底物检测CAT活性使检测过程简化，灵敏度也大幅度提高，具体应用可参照各公司产品说明书。

来自萤火虫的luciferase通过依赖ATP的方式氧化荧光素而发光，其最大峰值为560nm，在底物过量的条件下，光强与酶浓度成正比。萤火虫luciferase由*luc*基因编码，分子量为62000，目前被广泛用来作为植物、细菌、哺乳动物细胞及昆虫细胞的基因表达。由于发光检测背景低，使得*luc*作为报告基因检测极低水平的基因表达非常理想。荧光素酶的底物是荧光素，后者首先由Bitler和McElroy分离（从15000个萤火虫中提取了近9 mg）。目前应用的荧光素基本为人工合成的荧光素钠或钾盐（如图8-6所示），其生理作用与天然荧光素一致。用CCD影像的方法可以检测到单个细胞中的荧光素酶的表达。luciferase目前也可由重组技术生产，可广泛用于各种体外检测。

（四）生物素与亲和素系统

生物素-亲和素系统（biotin-avidin system，BAS）是20世纪70年代发展起来的一种新型生物反应放大系统。由于它具有生物素与亲和素之间高度亲和力及多级放大效应，并与荧光素、酶、同位素等免疫标记技术有机地结合，使各种示踪免疫分析的特异性和灵敏度进一步提高。BAS已经广泛应用于生物医学实验研究的各个领域，既可用于微量抗原、抗体及受体的定量、定性检测及定位观察研究，亦可制成亲和介质用于上述各类反应体系中反应物的分离、纯化。

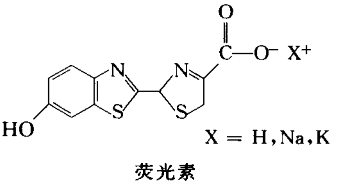


图8-6　荧光素的结构

1. 生物素（biotin，B）

生物素广泛分布于动、植物组织中，常从含量较高的卵黄（α型）和肝组织（β型）提取，α、β两型的生物活性基本相同，现已可人工合成。生物素在机体内以辅酶形式参与各种羧化酶反应，故又称为辅酶R或维生素H，分子量为244.31，分子中有两个环状结构，其中Ⅰ环为咪唑酮环，是与亲和素结合的主要部位；Ⅱ环为噻唑环，上有一戊酸侧链，其末端羧基是结合抗体和其他生物大分子的唯一结构。

利用生物素的羧基加以化学修饰可制成各种活性基团的衍生物，称为活化生物素，以适合与各种生物大分子结合的需要。主要有用于标记蛋白质氨基的生物素*N*-羟基丁二酰亚胺酯（BNHS）和生物素对硝基酚酯（pBNP），其中以BNHS最常用。近年来，应用活化长臂生物素（BCNHS）标记生物大分子，可以减少位阻效应，增加检测的灵敏度和特异性。用于标记蛋白质醛基、巯基和糖基的衍生物有生物素酰肼（BHZ）及肼化生物素（BCHZ）。

2. 亲和素（avidin，A）

亲和素亦称抗生物素蛋白、卵白素，是从卵白蛋白中提取的一种碱性糖蛋白，等电点

(pI)为10～10.5，含糖约10%，分子量为68000；纯品为白色粉末，易溶于水，在pH9～13的溶液中稳定，耐热（80℃加热2min）并耐受多种蛋白水解酶的作用，特别是和生物素结合后十分稳定；但对强光和Fe^{2+}比较敏感。亲和素由4个相同的亚基组成，能结合4个分子的生物素。亲和素与生物素之间的亲和力极强，二者结合的亲和常数（K_a）为$10^{15}mol^{-1}$，比抗原与抗体的亲和力（$K_a=10^{5\sim11}mol^{-1}$）至少高1万倍，因此二者能快速结合，而且反应不受外界干扰，具有高度特异性和稳定性。

3. 链霉亲和素（streptavidin，SA）

SA是链霉菌在培养过程中分泌的一种蛋白质产物。1L培养液中约含SA 10～60mg，主要通过α亚氨基生物素琼脂糖凝胶亲和层析法提纯。SA的分子量为65000，由4条序列相同的肽链组成。如同亲和素一样，一个SA分子也能结合4个生物素分子，二者亲和常数亦为$10^{15}mol^{-1}$。SA的每条肽链含159个氨基酸残基，其中赖氨酸、精氨酸等碱性氨基酸的含量比亲和素少，而酸性氨基酸含量较多，因此SA是一种稍偏酸性的蛋白质，其pI为6.0，并且不带任何糖基。

近年来，SA的克隆基因在大肠杆菌中的表达已获得成功。Sano等（1990）将构建的重组质粒转染细菌后，在加入适量外源生物素维持细菌生长的条件下，诱导5h，SA基因的表达蛋白占整个菌体蛋白的35%以上，用6mol/L盐酸胍透析（pH 1.5），可以获得近似均质的纯化SA，产量为3.9～6.5mg/100ml培养液，且活性良好。

三、利用抗体的标记技术

利用抗体的标记检测技术也称为免疫标记技术。它通常是用荧光素、放射性同位素、酶、铁蛋白、胶体金及化学（或生物）发光剂等作为追踪物，标记抗体或抗原进行的抗原抗体反应；并借助于荧光显微镜、射线测量仪、酶标检测仪、电子显微镜和发光免疫测定仪等精密仪器，对实验结果直接镜检观察或进行自动化测定，可以在细胞、亚细胞、超微结构及分子水平上，对抗原抗体反应进行定性和定位研究；或应用各种液相和固相免疫分析方法，对体液中的半抗原、抗原或抗体进行定性和定量测定。因此，免疫标记技术在敏感性、特异性、精确性及应用范围等方面远远超过一般免疫血清学方法。

根据实验中所用标记物的种类和检测方法不同，免疫标记技术分为免疫荧光技术、放射免疫技术、免疫酶技术、免疫电镜技术、免疫胶体金技术和免疫发光测定。

（一）免疫荧光标记技术

免疫荧光分析（immunofluorescence assay，IFA）始创于20世纪40年代初，1942年Coons等多次报道用异氰酸荧光素标记抗体。检查小鼠组织切片中的可溶性肺炎球菌多糖抗原，当时由于此种荧光素标记物的性能较差，未能推广使用。直至20世纪50年代末期，Riggs等合成性能较为优良的异硫氰酸荧光黄，Mashall等对荧光抗体的标记方法又进行改进，从而使免疫荧光技术逐渐推广应用。

1. 基本原理

免疫荧光技术是将抗原-抗体反应的特异性和敏感性与显微示踪的精确性相结合。以荧光素作为标记物，与已知的抗体（或抗原）结合、但不影响其免疫学特性。然后，将荧光素标记的抗体作为标准试剂，用于检测和鉴定未知的抗原。在荧光显微镜下，可以直接观察呈现特异荧光的抗原-抗体复合物及其存在部位。在实际工作中，由于用荧光素标记抗体检查抗原的方法较为常用，所以一般通称为荧光抗体技术。

目前用于标记抗体的荧光素主要有异硫氰酸荧光素（FITC）、四乙基罗丹明及四甲基异硫氰酸罗丹明。

2. 荧光抗体染色方法

（1）直接法　荧光抗体技术最简单和基本的方法。滴加荧光抗体于待检标本片上，经反应和洗涤后在荧光显微镜下观察。标本中如有相应抗原存在，即与荧光抗体特异结合，在镜下可见有荧光的抗原-抗体复合物。此法的优点是简单、特异，缺点是检查每种抗原均需制备相应的特异性荧光抗体，且敏感性低于间接法。

（2）间接法　根据抗球蛋白试验的原理，用荧光素标记抗球蛋白抗体（简称标记抗抗体）的方法。检测过程分为两步：首先，将待测抗体（第一抗体，简称“一抗”）加在含有已知抗原的标本片上孵育一定时间，洗去未结合的抗体；第二，滴加标记的二级抗体（简称“二抗”，是针对第一抗体的抗体）。如果第一步中的抗原抗体已发生结合，此时加入的标记二抗就和已固定在抗原上的一抗分子结合，形成抗原-抗体-标记二抗复合物，并显示特异荧光。此法的优点是敏感性高于直接法，而且只需制备一种荧光素标记的抗球蛋白抗体，就可用于检测同种动物的多种抗原抗体系统；缺点是有一定的非特异性荧光。此法常用于各种自身抗体的检测。如果第一抗体为已知，间接法也可用于鉴定未知抗原。

（二）放射免疫标记技术

放射免疫标记技术是将同位素分析的高灵敏度与抗原-抗体反应的特异性相结合，以放射性同位素作为示踪物的标记免疫测定方法。由于此项技术具有灵敏度高［可检测出皮克（pg）级微量物质］、特异性强（可分辨结构类似的抗原）、重复性好、样品及试剂用量少、测定方法易规范化和自动化等多个优点，因此在医学及生物学科的研究领域中广泛应用于各种微量蛋白质、激素、小分子药物及肿瘤标志物等的分析与定量测定。

1. 放射免疫测定

放射免疫测定（radio immunoassay，RIA）是 1959 年 Yalow 和 Berson 首先创建的经典放射免疫分析技术，用于血清中胰岛素含量的测定。30 多年来，由于此项技术灵敏、特异并已制成多种标准试剂盒，使用方便，应用范围十分广泛。目前，国外已成功地应用 RIA 检测的物质多达 300 余种，是测定各种微量物质不可缺少的手段。

基本原理：RIA 是标记抗原和未标记抗原对有限量抗体的竞争性结合或竞争性抑制反应。在 RIA 反应系统中，标记抗原（Ag*）、未标记抗原（Ag）和特异性抗体（Ab）三者同时存在时，由于两种抗原具有相同的决定簇，互相竞争结合抗体的能力相同，结果形成 Ag*-Ab 和 Ag-Ab 复合物。当 Ag* 和 Ab 的量固定时，二者结合形成免疫复合物就受到 Ag 含量的制约。如反应系统中 Ag 含量高时，对 Ab 的竞争结合能力就强，Ag-Ab 复合物的形成量就增加，Ag*-Ab 复合物则相对减少；反之，当 Ag 含量低时，对 Ab 的竞争结合能力弱，Ag*-Ab 复合物的形成量即增多。因此，Ag*-Ab 复合物的形成量与 Ag 含量之间呈一定的负相关函数关系。

2. 免疫放射测定（IRMA）

1968 年 Miles 和 Hales 应用同位素标记的抗胰岛素抗体检测牛血清中胰岛素获得成功，为了区别于经典的放射免疫测定（RIA），他们将其称为免疫放射测定或免疫放射度量分析（IRMA）。其基本原理是：待测抗原与过量标记抗体的非竞争结合反应，然后加入固相的抗原免疫吸附剂以结合游离的标记抗体，离心除去沉淀，测定上清液中放射性强度。从而推算出检品中抗原含量。由于在反应系统中使用过量的标记抗体，且无竞争性抑制反应，因此抗

体与待测抗原达到结合状态的化学平衡，在 2～3h 即可完成，较少受到抗体亲和常数的限制，即使单克隆抗体的亲和力较低，也能满足试验要求。同时一个抗原分子可以结合多个标记抗体分子，使 IRMA 的灵敏度明显高于 RIA。

3. 放射受体分析（RRA）

受体是存在于细胞膜表面、细胞浆或细胞核内的生物活性物质，其功能是和细胞外的信息分子配体特异性结合，然后将信息转变为生物效应。受体现已可从细胞或组织中分离、提取，进行定量和定位分析。放射受体分析或受体的放射配体结合分析，是建立在放射性标记配体与受体之间的结合反应，它是目前对受体分子进行定量和定位分析研究的灵敏、可靠的一项技术。

基本原理：放射受体分析的原理与免疫放射测定相似。应用放射性同位素标记配体，在一定条件下与相应受体结合成配体-受体复合物。由于二者的结合是表达配体与受体间的生物活性而非免疫活性，因此具有更高的特异性。放射受体分析可用于测定受体的亲和常数、解离常数、受体结合数以及定位分析等。

（三）酶免疫标记技术

免疫酶技术就是将抗原和抗体的免疫反应和酶的催化反应相结合而建立的一种新技术。酶与抗体或抗原结合后，既不改变抗体与抗原的免疫学反应的特异性，也不影响酶本身的酶学活性，即在相应而合适的作用底物参与下，使基质水解而呈色，或使供氢体由无色的还原型变为有色的氧化型（图 8-7）。这种有色产物可用肉眼、光学显微镜和电子显微镜观察，也可以用分光光度计加以测定。呈色反应显示了酶的存在，从而证明发生了相应的免疫反应。所以，这是一种特异而敏感的技术，可以在细胞或亚细胞水平上示踪抗原或抗体的所在部位，或在微克甚至纳克水平进行定量。

免疫酶技术的原理主要利用了酶催化作用，即很少量的酶即可导致大量的催化过程，所以极为敏感。它的催化过程有以下两种基本形式。

$$E+S \longrightarrow (ES) \longrightarrow E+P$$

$$E+S \longrightarrow (ES), (ES)+D1 \longrightarrow E+P+D2$$

E 为酶，S 为酶作用的底物，P 为底物分解后的产物，D1 为供氢体，D2 为 D1 的氧化型。如 P 或 D2 为有色化合物，即可用呈色反应显示酶的存在。

到目前为止，所应用的酶大多是辣根过氧化物酶，其次有碱性磷酸酶、酸性磷酸酶、葡萄糖氧化酶、β-D-半乳糖酶，每种酶通过与自己的特殊作用底物反应，而产生不同的颜色。

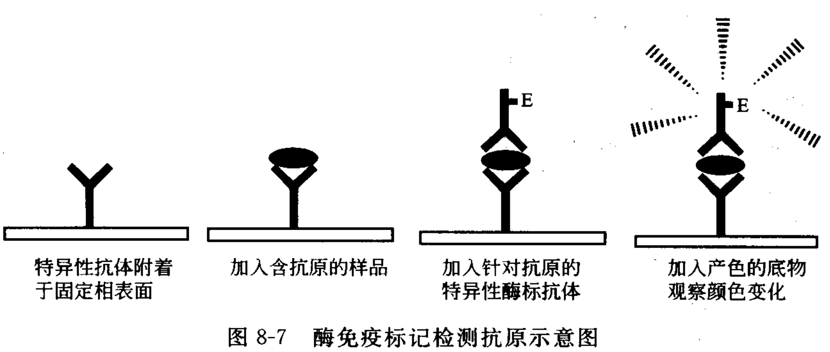


图 8-7 酶免疫标记检测抗原示意图

（四）胶体金标记技术

胶体金（colloidal gold）是氯金酸（chloroauric acid）的水溶胶，具有高电子密度，其颗粒直径在 1～30nm 之间，通过静电作用与多种生物大分子结合，已成为继荧光素、放射

性同位素和酶之后，在免疫标记技术中较常用的一种非放射性示踪剂。1971 年，Faulk 和 Taylor 首先报道将胶体金与抗体结合，应用于电镜水平的免疫细胞化学研究。1974 年，Romano 等用胶体金标记抗球蛋白抗体，建立间接免疫金染色法。此后，胶体金作为一种新型免疫标记技术发展很快。1978 年，Geobegan 等将胶体金标记抗体用于普通光镜下检测 B 淋巴细胞表面膜免疫球蛋白，建立了光镜水平的免疫金染色（immunogold staining，IGS）。1981 年，Danscher 用银显影方法增强金颗粒的可见度，并提高了灵敏度。1983 年，Holgate 等又加以改进，将免疫金染色与银显影技术相结合，增强检测信号的强度，使检测更灵敏，称为免疫金银染色法（immunogold silver staining，IGSS）。由于 IGS 和 IGSS 两种方法具有试剂制备简便、特异性强、灵敏度高、应用范围广并可用于双重和多重标记等优点，被广泛应用于医学和细胞生物学研究领域，是目前免疫组织化学技术中较常用的方法之一。近年来，胶体金标记技术进一步发展应用于免疫转印、流式细胞术（flow cytometry，FCM）、液相免疫测定、固相斑点金/银染色（dot IGS/IGSS）及斑点金免疫渗滤测定法（dot immunogold filtration assay，DIGFA）等多种标记免疫检测方法。

胶体金标记技术（immunogold labeling technique）是以胶体金作为示踪标记物，应用于抗原抗体反应的一种新型免疫标记技术（图 8-8）；胶体金是由氯金酸在还原剂如白磷、抗坏血酸、枸橼酸钠和鞣酸等作用下，聚合成特定大小的金颗粒，并由于静电作用成为一种稳定的胶体状态，故称为胶体金。由于它在中性及碱性环境中带负电荷，能与蛋白质分子的正电荷基团借静电吸引而形成牢固结合，从而标记抗体等蛋白质。此外，胶体金还可与其他多种生物大分子结合。根据胶体金的一些物理特性，如高电子密度、颗粒大小、形状及颜色反应，加上结合物所具有的免疫生物学特性，使其在免疫学、组织学、病理学及细胞生物学标记研究工作中得到广泛应用。

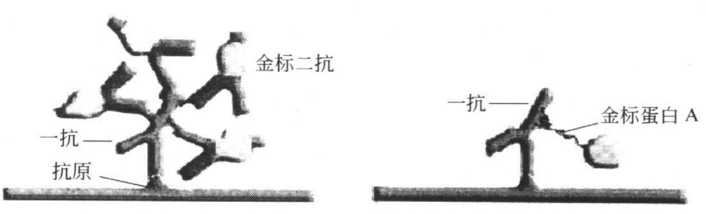


图 8-8　金标记抗体检测示意图

（五）发光免疫标记技术

发光免疫分析是将发光分析和免疫反应相结合而建立的一种新型超微量分析技术（图 8-9）。这种方法兼具有发光分析的高灵敏性和抗原抗体反应的高特异性。根据发光免疫分析中发光反应的不同体系和标记物及标记方法的不同，大体上可分为以下三种类型。

（1）化学发光免疫分析　以化学发光物质标记抗原或抗体，免疫反应后，直接引发化学发光反应进行检测。

（2）化学发光酶免疫分析　用参与某些发光反应的酶来标记抗原或抗体。免疫反应后，加入发光试剂，测定发光体系的发光强

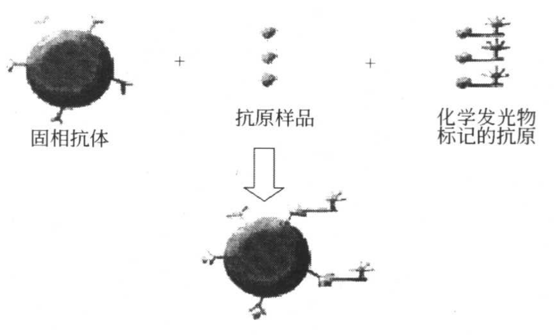


图 8-9　发光免疫标记示意图

度，进行抗原或抗体测定。

(3) 生物发光免疫分析　利用生物发光物质或参与生物发光反应的辅助因子标记抗原或抗体。免疫反应后，运用生物发光反应进行检测。

自从 Schroder 和 Halman 在 20 世纪 70 年代末期分别用化学发光免疫分析（chemiluminescence immunoassay，CLIA）测定甲状腺素（T4）以来，发光免疫分析技术发展很快。尤其是近年来，由于氨基酞肼类及其衍生物对抗原（或抗体）标记方法的建立，吖啶酯类等发光剂的合成，以及某些酶技术、超弱光检测技术的发展，进一步推动了发光免疫技术的进展，使之成为医学、生物学研究领域中一种新的手段。

（刘志国）

第九章　分子改造技术

本章所指的分子改造主要是指对大分子的改造，其目标主要针对蛋白质分子，因此从分类的角度来说属于蛋白质工程的范畴。通过分子设计和由设计结果所指导的特定的基因修饰，实现对天然蛋白质的定向改造，其目的是构造并生产特定功能的蛋白质新品种以及用于研究蛋白质的结构与功能关系等。蛋白质改造的方法并不是一个个地改变已存在的天然蛋白质，而是改造它们的基因，以改造后的基因作为模板，再用基因工程的方法来表达这些蛋白质产物或由工程菌来生产这些改造过的蛋白质。近年来，随着药物设计发展的需要，基于生物大分子结构知识的药物设计已成为药物发展中的重要领域，由于这种设计途径在很大程度上与蛋白质工程有着深刻和广泛的联系，因此往往在学术上与蛋白质工程一起讨论。

第一节　概　　述

近年来，随着一些新技术的发展，蛋白质的结构数据日益增多，以及基于计算机技术及各种物理学原理的模拟方法的发展，为分子改造提供了良好的基础。但是，应该看到对蛋白质的改造是一个复杂的、多学科交叉的研究领域。而且改造也不是一次就能够完成，往往要经过多次的循环，不断改进设计方案才能发现一个理想的改性蛋白。由于蛋白质改造巨大的应用潜力，使其成为近年来快速发展的领域之一。

1. 发展简史

蛋白质改造技术的应用历史并不长，其发展是和一些科学家开创性的工作分不开的。早在 1987 年美国杜邦公司的 DeGrado 等经过精心的设计并由多肽合成法产生出了一个由 74 个氨基酸组成的非天然蛋白质。它具有设计者所期望的典型的 4 股 α 螺旋结构。这一工作被称为蛋白质分子设计的第一个里程碑。随后，美国 Duke 大学的 Richardson 等设计并用基因工程的方法表达产生了另一种 4 股螺旋的非天然蛋白质，同时还表达产生了全 β 折叠结构的、称为 betabellin 的非天然蛋白质。此后，对改造方法进行了不断地研究与完善。一系列的实验说明，定向地设计与表达产生具有人们所要求的空间结构与性质的蛋白质是合理的、可行的。蛋白质的分子改造在药物设计、研究与生产中蓬勃地开展起来，取得了许多成功的例子，如艾滋病毒 HIV 蛋白水解酶抑制剂的设计与合成、胰岛素的改造、抗体的改造等。

2. 蛋白质改造的对象

蛋白质改造研究目前主要集中在两个方面：一方面是对具有明显生物学功能的天然蛋白质分子，像胰岛素、天花粉及抗体等的改性；另一方面是基于生物大分子结构知识的药物设计与表达生产。传统的药物研制主要是从大量的天然产物如动物、植物、微生物，以及合成有机、无机化合物以及矿物中进行筛选。往往得到一个可供临床使用的药物要筛选 1 万种不同的化合物，要经过 10 年左右的时间和耗资 2.5 亿～3.0 亿美元。利用结构生物学的发展

所获得的大量的蛋白质、核酸、糖类的三维结构数据，设计合成相关药物，则可大大降低研究成本与开发周期。这是因为任何药物都是作用在相应的受体上的，这些受体或靶分子往往又是蛋白质或核酸，如果有了这些大分子的结构，就容易发现哪些物质能与这些受体发生特定的结合，使得药物的筛选与设计有了科学依据，避免了盲目性。

原则上，任何有功能的蛋白质都可以作为改造对象，但实际上选择目标时往往要考虑以下几个方面：①改性对象有没有测定出空间结构。如果没有测定出空间结构，有没有同源蛋白的空间结构，如果都没有，那就很难进行以后的分子设计与改造。②结构和生物功能的联系是否明确。功能尚不清楚的蛋白质对象改造起来有一定的盲目性。应当先确定其功能部位或分区，然后确定改造的方向。③改造的实际意义。这需要根据所选对象的重要性确定，即改造所能产生的科学价值或经济效益与社会效益。④可行性分析。改造后是否易于通过基因工程生产。上文提到的胰岛素、抗体等都是很好的蛋白质改造对象；此外，还有目前已经上市的多种基因工程药物，如凝血酶、白介素、肿瘤坏死因子等。

3. 蛋白质改造的基本方法

蛋白质改造的基本方法是通过人工方法制造突变，或称诱变。诱变主要分为两种：一种是定点诱变，另一种是随机诱变。前者包括点突变、片段的插入或缺失等方法，后者则常采用化学诱变、盒式诱变、错掺诱变和寡核苷酸掺入诱变等方法。在定点诱变中，需要利用各种生物化学、晶体学、光谱学等方法对天然蛋白质或其突变体进行研究，获得其分子特征、空间结构、结构和功能之间关系以及氨基酸残基功能等方面的信息，并以此为依据对分子进行改造，故称为分子的合理设计；与此相对应，不需要准确的蛋白质分子结构信息而通过随机突变和筛选等方法对其进行改造，则称为分子的非合理设计。两者均具有很强的实用性，并且各有特色。前者是在已知结构、功能信息基础上的直接改造；而后者则利用了自然进化机制（随机突变、基因重组和自然选择），在体外人为地创造特殊的进化条件，模拟自然进化过程，筛选获得期望的分子，即所谓定向分子进化。

第二节　蛋白质的体外改造

蛋白质的改造主要采用体外诱变方法。通过体外诱变改变编码蛋白质的 DNA 片段的碱基序列。这些改变可以是局部的或普遍的，随机的或靶向的。普遍的和非特异性的诱变方法更适用于分析基因的调控区，而精确的诱变则常用于推断单个氨基酸或多个氨基酸对靶蛋白的结构和功能的作用。两种诱变方法都具有在体外产生诱变体而不需进行表型筛选的优点。

从 20 世纪 80 年代初期开始，蛋白质诱变修饰已经成为常规方法，包括寡核苷酸指导的位点特异性诱变、从寡核苷酸合成全基因、扫描诱变、随机诱变等。

一、寡核苷酸指导的诱变方法

这种方法的具体做法很多，文献也非常丰富，但原理基本一致。它一般包括 3 个阶段：①设计 1 个或多个寡核苷酸，并与待突变单链 DNA（ssDNA）的一定区域退火；②在 DNA 聚合酶指导下沿寡核苷酸链延伸合成具突变位点的 DNA 链；③转化后，筛选获得突变

克隆。

（一）非 PCR 的诱变方法

在非 PCR 的诱变过程中，将待修饰的 DNA 序列克隆到适当的载体中（如质粒、噬菌体等），允许在体内扩增待修饰 DNA 序列。精确设计带有突变序列的寡核苷酸链，它与单链 DNA 模板退火，如构建在单链噬菌体载体（如 M13 或 Φ174）中的形式。退火的寡核苷酸作为 DNA 聚合酶（通常是 T4 或 T7 DNA 聚合酶）体外合成 DNA 的引物，在 DNA 连接酶存在的条件下，产生闭合的环状双链 DNA 分子。将形成的双链 DNA 引入合适的宿主细胞进行复制，由此导致亲代链和突变链的分离，然后通过大量的筛选获得所要求的序列修饰。上述过程如图 9-1 所示。

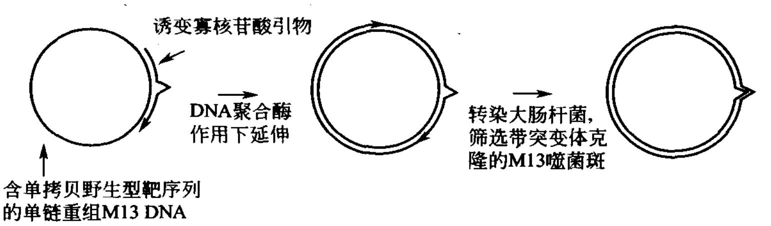


图 9-1 单链 DNA 的非 PCR 突变策略

对于单点突变，合成 15～22 个核苷酸长的寡核苷酸链，将待突变的序列置于链中心，以过量物质的量比与单链 DNA 混合、退火。对于更复杂突变，通常需要在突变点的两侧都有 15～18 个核苷酸，这将有利于与模板序列形成完全匹配。同时加入 DNA 聚合酶和 DNA 连接酶，单链 DNA 区转变为双链 DNA，并通过连接酶闭合起来。常使用 T7 DNA 聚合酶，它具有有效的 3′→5′校正活性和较高的连续合成能力（称为高续进性），但无 5′→3′核酸外切酶活性。突变体和亲代序列通过转化到一合适的宿主，并允许复制、发生分离；然后就能通过大量的筛选鉴定突变体序列。

上述经典的定点诱变方法中，含有预定突变的克隆比例为 0.1%至 50%不等，这取决于诱变寡核苷酸的效率。现在，发展了几种筛选突变 DNA 克隆的方法（或筛除含非突变 DNA 克隆的方法），包括：①通过限制性酶，或外切核酸酶等去除修饰碱基，选择性消化模板 DNA；②恢复抗生素抗性基因或复制起始子；③从突变体 DNA 中删除限制位点。

经过多年的研究改进，上述寡核苷酸介导定点诱变的方法已极为有效和可靠。然而，PCR 技术的诞生，使这种定点诱变的经典方法很快被弃用，它们已不再是 DNA 靶片段产生突变的首选方法。近来发展的许多 PCR 方法使得绝大多数定点诱变比经典方法更加容易。

（二）PCR 方法

目前使用的绝大多数以 PCR 为基础的诱变方法是热稳定 DNA 聚合酶发现以后，在 PCR 技术上衍生而来。由 PCR 扩增的原理知道在寡核苷酸引物与靶 DNA 杂合分子中央位置的单个碱基错配并不影响扩增效率。而这些错配被证明能够将模板 DNA 序列转换为 PCR 产物中引物区域的定点诱变序列，换句话说，能够替换同源的野生片段。这种诱变方法的主要优点是：①突变体的回收率很高，以致无需排除针对非突变模板的筛除；②能用双链 DNA 为模板，并几乎可在任何位点引入突变；③使用较高的延伸温度，以降低模板 DNA 形成二级结构的能力，提高延伸反应效率；④所有的反应可在同一只试管中进行，同时也无

需在噬菌体 M13 载体上克隆，使方法及操作快速简便。当然以 PCR 为基础方法存在缺点，较为突出的是 PCR 产物的错误率相对较高，即除预定突变外，常包含一些非预定突变。然而这一问题并不难解决，可通过减少扩增的循环次数及应用高保真的耐热 DNA 聚合酶等方法降低错误率。正因为如此，近来基于 PCR 方法的应用越来越广泛。

在已报道的以 PCR 为基础的各种诱变方法中，有两种方法的实用性非常突出：重叠延伸诱变和大引物诱变。

1. 重叠延伸诱变

在重叠延伸诱变中，两个重叠的 DNA 片段是分别从两个独立的 PCR 扩增反应得到的。突变位点设计在重叠区域，并存于两个扩增片段中。将重叠片段混合后，用两条分别与两个原始模板末端结合的引物进行第三次 PCR 扩增，获得全长 DNA，这种方式具有惊人的效率，但需两个诱变引物，两个侧翼寡核苷酸引物，三次 PCR 反应才能构建突变，如图 9-2 所示。

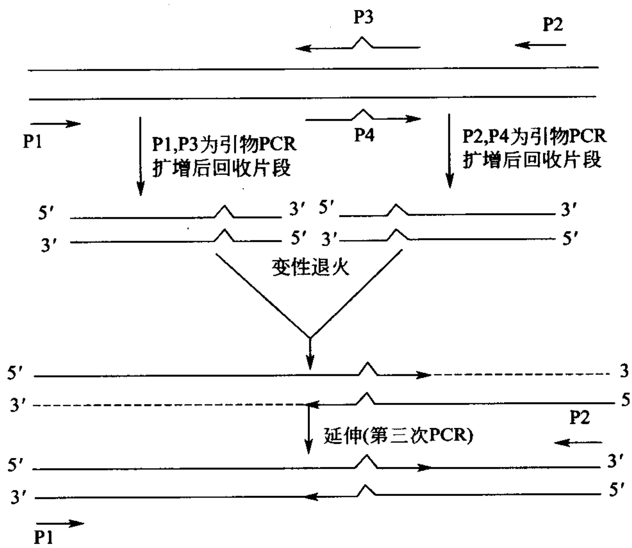


图 9-2 重叠延伸法构建突变体

P1、P2 为两侧引物，P3、P4 为突变引物且两者碱基互补。突变碱基特别设计在引物的中部（尖角符号表示）。以 P1、P3 为引物 PCR 扩增左侧片段；以 P2、P4 为引物 PCR 扩增右侧片段。左、右两侧片段回收纯化后混合，经变性退火后即可产生部分重叠片段，并在随后的反应中延伸形成突变体。在这一反应中加入外侧引物（P1、P2），融合突变体一旦形成便可呈指数扩增。最终产物可利用序列内的位点或在引物 P1、P2 内被引入的位点，将该片段克隆至合适的载体

某些情况下，如果能够策略地把限制酶切位点设置在含有突变的扩增 DNA 片段上，并进行克隆，可使用一个较为简单的变通方法（一个诱变引物，两个连续的 PCR 反应）完成重叠延伸诱变反应。当然，并非所有的反应都能采用这种方式。

如果需要改变的片段较长，不能设计在突变引物中，则可采用片段缺失诱变的方法。在这种方法中，同样包括 PCR 扩增上游（左侧）片段和下游（右侧）片段的步骤，也要进行回收片段后混合及用两侧引物进行 PCR 扩增反应。其主要区别在于突变引物的设计，此时的突变引物 5′端与模板之间并不互补，但必须与另一突变引物的 3′端互补。这样，在 PCR 扩增的上、下游片段回收、混合后，在变性退火条件下两片段能借此互补区结合并进行延伸

反应，从而能够扩增出去掉了缺失区的 DNA 片段。上述过程如图 9-3 所示。

上述方法也可扩展用于将来自不同蛋白的结构域的独立的编码序列连接在一起，或者允许一个蛋白内的结构域重组。此时扩增的模板不一定是同一条链，但引物的设计原则是相同的。应当注意此方法中突变引物的设计需非常精确，否则可能导致密码阅读框的改变或其他不正确的拼接。

除了上述两种方法以外，目前还有许多类似的方法，其原理或思路基本相似，在此不一一列举。

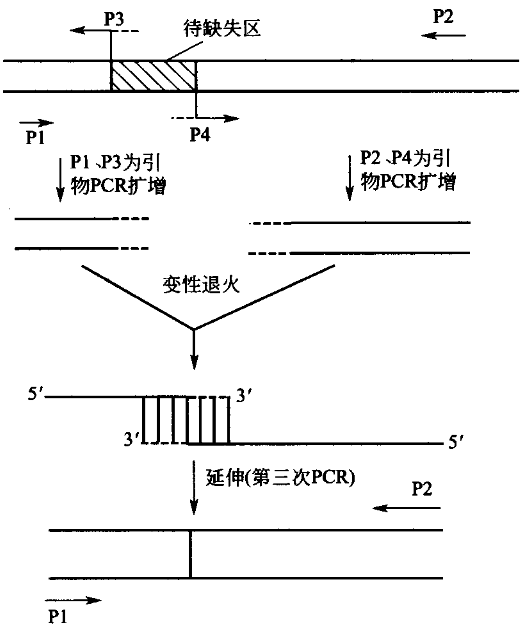


图 9-3　较大片段的缺失诱变

P1、P2 为两外侧引物。P3、P4 为突变引物且两者碱基互补，但它们的 5′端（图中虚线部分）与待缺失区的序列不互补。同样，以 P1、P3 为引物 PCR 扩增上游（图中左侧）片段；以 P2、P4 为引物 PCR 扩增下游（图中右侧）片段。左、右两侧片段回收纯化后混合，经变性退火后即可产生部分重叠片段，并在随后的反应中延伸形成突变体。以此延伸产物作为模板，用外侧引物（P1、P2）进行 PCR 扩增即可获得突变产物

2. 大引物诱变

大引物诱变方法是 Kammann 等（1989）提出的，随后经改进完善。它是目前以 PCR 为基础的诱变方法中最为简单、经济的方法。该方法包括两轮 PCR，需两个外侧引物，一个含碱基替换的诱变引物。外侧引物能够与克隆基因或邻近载体序列互补。理论上，诱变引物的扩增方向可以朝向任何一侧。然而，实际上诱变引物总是朝向较近一侧的引物，这样能使反应产生的大引物的长度保持最短。

首先，以野生型 DNA 为模板，用诱变引物和较近一侧的外侧引物进行第一轮 PCR 扩增。第一轮反应的产物（双链大引物）经纯化后与另一侧引物进行第二轮 PCR。所得产物为包含突变的双链 DNA，大小为两个外侧引物间的距离。

大引物方法的主要问题是残留的突变引物对第二轮 PCR 扩增的干扰。因此需要花费精力纯化大引物（常用琼脂糖凝胶电泳回收纯化步骤）。一个巧妙且简单的解决方案是用具有显著不同解链温度的外侧引物，去引导两轮 PCR。第一条外侧引物采用 15 或 16 个碱基的长度，T_m 值为 45℃。因此，第一轮 PCR 是在低退火温度下进行的。在第一轮 PCR 后期，第二个外侧引物（25～30 个碱基长度，退火温度为 72～80℃）直接加到反应管中，用高退火温度（通常 72～80℃）进行 PCR 扩增。最后由大引物和第二个外侧引物合成全长 DNA 产物。后期扩增反应中，由于退火温度高，第一个外侧引物的退火被抑制，因此，第二轮 PCR 的产物几乎全是突变 DNA。大引物诱变方法如图 9-4 所示。

（三）诱变寡核苷酸的设计

定点诱变的关键步骤是设计诱变寡核苷酸。通常，诱变寡核苷酸必须包含至少一个碱基的改变，也可以包含更复杂的突变，例如插入、缺失及复合物替换等。诱变寡核苷酸的最低长度是由突变的复杂程度决定的。用 25 个碱基长度就可完成简单的单碱基替换。较复杂的突变则需要更长的碱基。设计中应综合考虑到其他性质，如解链温度、形成二级结构的倾

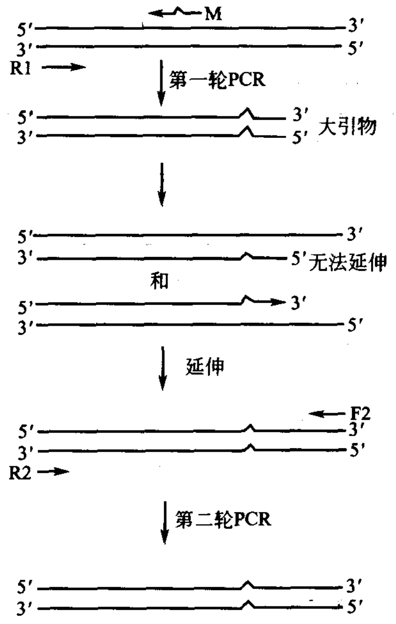


图 9-4　大引物诱变方法
在第一轮 PCR 中，用正向诱变引物（M）和反向引物（R1）扩增出双链大引物。第二轮 PCR 的第一循环过程中，大引物以原始 DNA 为模板引导链延伸。以后由正向引物 F2 和反向引物 R2 进一步扩增出希望的突变体

向、结合的特异性等，以最大限度地保证有效的诱变。

传统（非 PCR 方法）的定点诱变寡核苷酸必须符合以下要求。

① 与靶 DNA 的适当部位互补。

② 与靶序列特异结合的碱基不低于 20 个。2 个或更多的碱基替换时，寡核苷酸的长度至少需要 25 个碱基。

③ 携带的错配碱基应位于中央位置。这样每侧有 10～15 个碱基可与模板链完全匹配。超过 3 个核苷酸的插入、缺失或替换时，要求引物在诱变点的每侧有 30 个核苷酸互补于模板，且引物两侧与靶 DNA 互补形成双链的热稳定性应该尽可能相等。

④ 无回文、重复或自身互补序列，否则寡核苷酸与模板 DNA 靶序列的杂交率很低。

以 PCR 为基础的定点诱变方法的设计遵循同样的基本准则，但应注意以下几个问题。

① 使用 *Taq* 酶扩增时，可在扩增产物的 3′端加上一个与模板序列无关的附加核苷酸（通常为 A），许多情况下，附加的核苷酸不与模板链配对。为此，应做适当调整以为恢复 PCR 产物与靶 DNA 间的一致性，或改用无此作用的其他耐热 DNA 聚合酶。

② 许多情况下，可在突变的寡核苷酸上引入新的酶切位点或消除靠近突变点的已存在的酶切位点。这种策略有利于后期用限制性内切酶消化来筛选候选克隆，鉴别突变体。

③ 设计引物时应仔细分析突变引物序列是否具有与模板的其他区域而非靶序列形成稳定杂交分子的能力，如发夹形成、二聚体形成、可能的错误退火位点、突变或插入的位置和数目、熔链温度、5′端和 3′端的稳定性等。很多引物设计程序都具有这些分析能力。

二、随机诱变

随机诱变有 4 种常用方法：化学诱变、盒式诱变、错掺诱变和寡核苷酸掺入诱变。

（一）化学诱变

化学诱变是局部随机诱变中较早采用的，也是目前最常用的方法之一。该法主要利用化学诱变剂处理 DNA，产生随机的诱变，然后用已突变的片段构建一个重组克隆库，通过功能分析鉴定产生新表型的重组子。突变的精确位置可通过测序等方法确定。化学诱变使用的诱变剂主要为亚硝酸、羟胺、亚硫酸氢盐或肼等。

化学诱变的主要缺陷是回收的突变体的频率非常低。而且，大多数通常所用化学诱变剂都以特异的方式同碱基作用，这样由单种诱变剂获得突变的范围可能太窄，不能广泛地分析蛋白质独特片段。最后，双链 DNA 的化学诱变是高度非随机的，倾向变性的区域最易受到攻击。以单链 DNA 作为诱变的底物比用双链情况好一些，能使这些问题减轻但不能完全消除。用化学试剂处理单链 DNA，损伤的碱基有一定选择性，如亚硝酸产生转换突变，即嘌呤替换嘌呤，嘧啶替换嘧啶；羟胺主要导致 C-G 碱基对转换为 A-T 碱基对；亚硫酸氢

盐几乎只能与胞嘧啶作用，脱氨后形成尿嘧啶，因此亚硫酸氢盐同羟胺产生同样类型的转换突变；肼能裂解胞嘧啶和胸腺嘧啶的杂环，使其水解。这些化学试剂修饰单链 DNA 上的碱基而不破坏磷酸二酯键骨架。除去化学物后，可用 DNA 聚合酶在引物引导下合成互补的 DNA 链。在合成过程中，当聚合酶遇到模板链中的损伤碱基时，完全随机地插入核苷酸。因为 4 种核苷酸都能在单个位置上插入，因此每个损伤点的突变概率为 75%。由于碱基颠换频率是碱基转换的两倍，导致突变产生的 DNA 具有相当大的改变。延伸反应完成后，产生的双链靶片段被克隆入适当的载体，可通过 DNA 序列分析直接鉴定突变克隆。

（二）丙氨酸扫描诱变

丙氨酸扫描诱变是改变蛋白质表面残基而不改变蛋白质三维空间结构的一种方法。该方法最初是用于发现与其他蛋白质或小配基相互作用的蛋白质表面残基。蛋白质中典型的带电荷残基有精氨酸残基、天冬氨酸残基、谷氨酸残基等。这些残基的带电荷侧链通常在蛋白质表面，暴露于溶剂中，它们同底物、抑制剂或其他配基（包括蛋白质）相互作用。通过定点诱变可消除蛋白质上的某些侧链，即用丙氨酸代替蛋白质表面的带电荷残基。这通常不干扰蛋白质核心区的折叠，但是关键部位的极性基团或带电基团的替换，将可能导致蛋白质功能的严重损害，从而确定特定功能所需的关键基团。

带电荷氨基酸的丙氨酸扫描诱变产生一套可通过功能丧失进行分析的有规则的突变蛋白群，而避免了一个大的随机突变体文库建立、测序和对其性质的了解。即使这样，由广泛的丙氨酸扫描产生的突变体的数量很大。为减少工作量，可将带电荷氨基酸归并成簇。每簇含有精氨酸残基、赖氨酸残基、天冬氨酸残基、谷氨酸残基和组氨酸残基，而在这 5 种残基中，其一级序列彼此相互成簇即可。在这些簇中，每簇的带电荷氨基酸由寡核苷酸定点诱变方法同时变为丙氨酸残基，然后进行表达及分析。应当注意在丙氨酸扫描中包含组氨酸残基的替换结果是不可靠的，因为由丙氨酸替换组氨酸可能因此使突变体蛋白质不稳定或导致其他不可预料的结构改变。

丙氨酸扫描不适于内部不稳定的蛋白质及分子较大的蛋白质，对中等大小蛋白质的表面上带电荷残基或带电荷簇的全面详细的研究，工作量也较大，需要建立许多突变体，并进行表达和分析。对较小的蛋白质或对单个结构域的功能残基研究或蛋白质亚结构的详细图谱研究，推断不同侧链在已知结构蛋白质中的稳定性和独特区域折叠中的作用研究等方面，丙氨酸扫描则是一个可选择的方法。

三、突变体筛选方法

寡核苷酸介导诱变的频率在 0.1%～50%，因此在克隆筛选过程中会有相当多的非突变的克隆或噬菌斑，有时数目相当大。当以双链 DNA 质粒为模板时，PCR 为基础的诱变也存在这个问题。

为克服这些困难，开发了许多体外选择性破坏非突变 DNA 或抑制野生型克隆生长的方法。应用这两种方法，定点诱变后恢复的重组子中，突变克隆的比例得到了增加。在引物延伸 PCR（如大片段缺失突变和多点突变）或用寡核着酸库同时在靶 DNA 的精确位置产生许多突变体和试图分离“困难的”的突变体时，就必须通过这些方法筛选突变的克隆。突变体筛选的方法很多，以下主要介绍 2 种常用的筛选方法。

（一）限制酶切割区分亲本模板与突变产物

这种方法利用了细菌的限制-修饰系统，可以选择限制酶，以切割亲本链而不破坏突变链的方法，如利用限制酶 *Dpn* Ⅰ特异性切割双链 DNA 中甲基化序列 G^{m6} ATC 的方法，使内源性甲基化的亲本链被切割，而用 4 种通用脱氧核糖核苷酸体外合成的 DNA 没有甲基化，因而完全抵抗切割。在定点突变后，能用 *Dpn* Ⅰ降解剩余的甲基化了的野生型模板，从而富集体外合成的未甲基化 DNA。由于 *Dpn* Ⅰ选择性地破坏亲本模板，在 PCR 为基础的诱变中，这一方法的主要用途是净化合成的双链 DNA。

另一种区别模板链的途径是在体外合成时将甲基化的三磷酸核苷酸掺入到突变链内。突变链将免受如 *Hha* Ⅰ、*Sau* 3A Ⅰ、*Msp* Ⅰ等限制性内切酶的消化。这些酶将把没有修饰的亲本模板 DNA 降解成小片段。

（二）单一限制位点的消除

在标准寡核苷酸定点诱变中，采用单一位点删除的目的是消化掉在非必需区域含有单一限制位点的未突变双链质粒 DNA。该方法中，两条引物同时与变性模板杂交，一条引物在靶 DNA 上引入预设突变，另一条引物破坏模板质粒上的单一限制位点。在体外，突变反应完成后，通常将质粒 DNA 转化大肠杆菌中。

该菌株是缺陷株（如 *mut*S^-突变体），不能纠正错配的碱基对。异源双链 DNA 中错配碱基对是通过复制恢复而非修复。从转化子库中提取质粒 DNA，用在独特位点裂解野生型 DNA 的限制酶消化，没有突变的质粒将被裂解为线性 DNA，而线性 DNA 难以转化入大肠杆菌。另一方面，突变体质粒抗消化，保持转化能力，在转化入 *mut*S^+ 菌株后，被有效地恢复。

若两个突变点放在一条引物上（可以从新合成或 PCR 方法产生），能增加该法的工作效率。在下面两种情况下，特别有用：①若限制位点为 *Bam*HⅠ、*Bgl*Ⅱ或 *Pvu*Ⅱ等，若用两条独立的诱变引物时，由于未知原因，基于这些酶识别位点的选择难以奏效；②模板质粒大于 6kb。

四、分子进化

（一）概念

在研究中有时并不清楚真正的目标氨基酸位于什么位置，要把它或它们改成什么。此时的策略是利用进化的思想，在体外通过随机的突变产生一系列突变体或突变体文库，通过定向的筛选、测试获得目标产物，这就是所谓的分子进化改造方法。这种方法依赖于三个主要特征。

第一，编码感兴趣蛋白质序列的核酸与蛋白质保持着物理联系。这种联系可以通过把蛋白质呈现在噬菌体、细菌或真核细胞表面。

第二，存在产生大量突变体的方法。最直接的方法是通过易错 PCR（epPCR）或饱和突变等把随机突变引入目的基因。采用这种方法，因为包含的突变组合的数量之多，一次只能改变一个蛋白质的一些小片段。文库受到宿主吸收个别成员的能力的限制，认为 10^{12}～10^{14} 是个很大的数量；然而，即使是在小蛋白质的结构域中，所有残基可能组合的数量也远远超出这个数量。

第三，这些方法需要一个筛选或选择策略来从文库中富集含有所需目的表型的新的蛋白质序列。选择系统通常包括宿主内的重要功能的互补或设计结合到配体的筛选方法。有人描

述了既有催化作用又可使呈现蛋白被捕捉到固相支持体上的筛选方法。也有一些新的体内选择系统，可在一个完整的机体内进行选择。

（二）基本方法

1. 易错 PCR

易错 PCR（error-prone PCR，epPCR）是指在利用 *Taq* 酶进行 PCR 扩增目的基因的同时引入碱基错配，导致目的基因随机突变。由于一次突变很难获得满意的结果，故可进行连续的易错 PCR 方法扩增，即将一次 PCR 扩增得到的有益突变基因作为下一次 PCR 扩增的模板，连续反复地进行随机突变，使每一次获得的小突变累积而产生重要的有益突变。

该方法属于无性进化方法，其基因改变只发生在单一分子内部。它虽原理简单，但操作较为费力、耗时，通常适用于 800bp 以内的较小基因片段。此外，使用该方法易出现同型碱基转换。

2. DNA 改组和族改组

DNA 改组（DNA shuffling）是通过改变蛋白质的基因，使蛋白质体外加速定向进化的方法。其基本步骤是将选定的单一基因用 DNase Ⅰ 消化，使之片段化，继而进行体外重组，经标准的克隆筛选步骤获得突变体克隆。上述步骤可重复，直到获得满意的结果（图 9-5）。当 DNA 改组用于一组同源基因时，这种技术称为“族改组”（family shuffling）。族改组的重组子的产率较低。当改组具有 80％核苷酸顺序同源性的两个同源基因时，杂合体形成的频率低于 1％，低的原因可能是由于异质复合体形成的频率低。为此，设计了两个改进的族 DNA 改组技术：①单股的族改组，即分别制备两个同源基因的单链 DNA，其中一条链是基因的编码链，另一条是另一基因的非编码链。两个单链 DNA 分子用 DNase Ⅰ 消化后，进行同源重组。②用限制酶消化 DNA 片段进行 DNA 的改组。限制酶消化的 DNA 片段退火可以高频率地产生同源复合体，但是只有异源复合体分子能进行 DNA 的延伸。

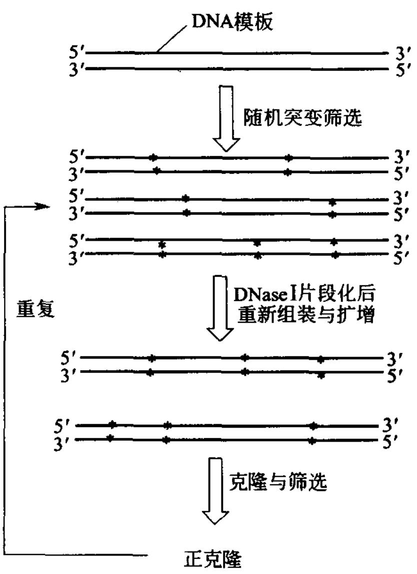


图 9-5 DNA 改组示意图

族改组是各种来源的同源基因重组广为应用的技术，可用于多种同源基因产生功能嵌合蛋白质库。例如，用 20 个不同来源的人干扰素-α 基因的改组产生嵌合库，它含有抗病毒和抗增殖活性都提高的各种突变体。其中高活性的嵌合体序列来自 6 个母体基因顺序片段，它比最好的母体抗病毒活性提高 87 倍。连续几轮 DNA 改组产生的嵌合体活性超过最好的人干扰素母体 185 倍。有趣的是，对某些进化性质而言，最好的母体与最好的嵌合体在顺序上不是紧密相关的，这表明最好的母体不一定是单一基因进化最好的起点。

3. 合成改组

合成改组（synthetic shuffling）是一种定向进化的新方法，母体的每个氨基酸可以同每个其他的氨基酸单独地组合。合成改组提供由数据库顺序到功能库的一条直接路线。15 种枯草杆菌蛋白酶基因的合成改组可获得不同性质组合的嵌合酶，用其他定向进化方法得不到

此结果。

片段改组的突变体都与一个或多个原始母体相关。相反，合成改组的突变体的特性倾向于彼此远离，不再簇集在母体周围。合成改组使得突变体多样性增加，是由于破坏母体基因所表达的蛋白质中氨基酸之间的组合，并降低突变体中氨基酸序列与母体中的相似性。许多合成改组产生的蛋白质 20%是有活性的，并具有新的功能。

合成改组使现有蛋白质的顺序空间扩大。合成改组在蛋白质库的构建上有很大的灵活性。由于所有的基因是合成的，因此可根据宿主对密码子的偏爱特性，选择合适的密码子以便在宿主中有效地表达。对于同源性较低的母体基因，交换体也可直接在氨基酸水平上合成产生。此外，合成的多样性不受母体基因的限制，这样，提供了由顺序数据库到蛋白质库的直接路线。在任何已知蛋白质中不存在的变化也可以通过合成改组的方法实现。因此合成改组是产生蛋白质顺序多样性和功能多样性最有效的方法，很少受母体基因之间顺序同源性的限制。

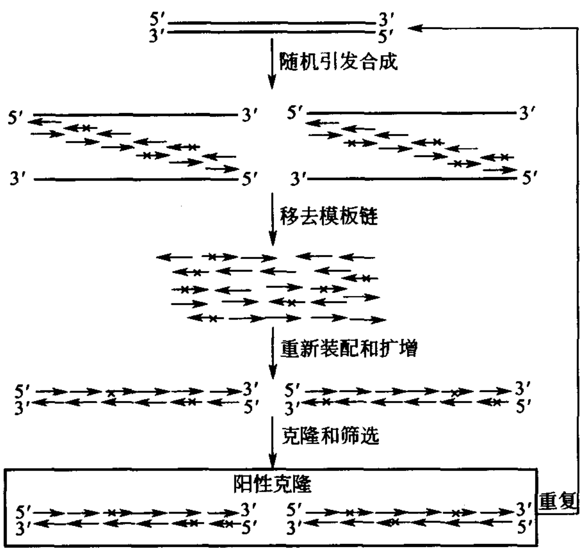


图 9-6　体外随机引发重组原理示意图

4. 体外随机引发重组

体外随机引发重组（random-priming in vitro recombination，RPR）是以单链 DNA 为模板，配合使用一套随机序列引物，产生大量互补于模板不同位点的短 DNA 片段，这些短片段在随后的 PCR 反应中，互为引物进行合成，产生各种组合片段，经克隆组装成完整的基因。如果需要，可反复进行上述过程，直到获得满意的进化产物（图 9-6）。上述过程中，由于碱基的错配等，也会在 DNA 片段中产生少量的点突变。RPR 可以利用单链 DNA 为模板进行 PCR 扩增，因此其对亲本 DNA 的需要量较低，同时随机引发的 DNA 合成不受 DNA 模板长度的限制。该方法中合成的随机引物长度相同，无顺序倾向性，在理论上可使模板上每个碱基以相似的频率发生突变。

5. 交错延伸

交错延伸是在 PCR 反应中把常规的退火和延伸合并为一步，并限定延伸时间（55℃，5s），因此只能合成出非常短的新生链。经变性的新生链再作为引物与体系内同时存在的不同模板退火而继续延伸。此过程反复进行，直到产生完整的基因长度，结果产生间隔的含不同模板序列的新生 DNA 分子（图 9-7）。这一过程采用的是变换模板机制，与逆转录病毒所采用的进化过程一致。

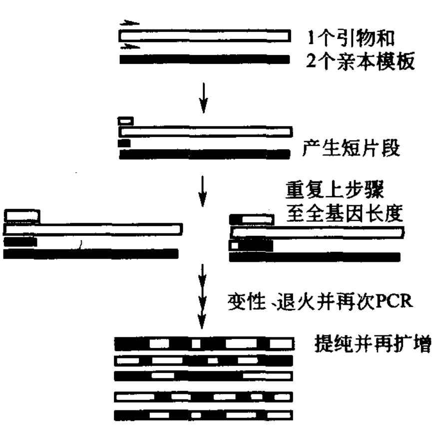


图 9-7　交错延伸原理示意图

交错延伸法重组发生在单一试管中，不需分离亲本 DNA 和产生的重组 DNA，因此方法简便有效，

是定向分子进化的又一强有力的工具。

第三节　蛋白质改造技术的应用

蛋白质改造技术应用于生物领域的各方面，作为产生新分子、理解蛋白质基本性质的重要方法。以下将介绍该方法的具体应用实例，包括最简单的点突变和结构域重组及更为复杂的多点改变等。

一、点突变

只要有了基因序列，使用前面提到的突变技术就很容易实现蛋白质中单独的点突变。已经知道大量的氨基酸序列变异体的例子，下面给出其中的一些例子。

一个最早的药物蛋白质改造的例子是干扰素 β-1b 的生产。干扰素 β 有 154 个氨基酸，其中第 17 位残基，用半胱氨酸（Cys）替代丝氨酸（Ser）后就产生这个新蛋白质——干扰素 β-1b。这个替代减少了在大肠杆菌中合成期间错误二硫键形成的可能性，同时也去除一个转录后氧化的可能位点。大肠杆菌中表达的干扰素 β-1b，其活性接近于天然 β-干扰素。

另一个例子是胰岛素的改造，成熟胰岛素由 A 链和 B 链组成，将其中 B 链 C 端的两个残基——第 28、第 29 位的脯氨酸和赖氨酸的顺序颠倒即产生了一个新的工程胰岛素 Humalog（优泌乐）。Humalog 是胰岛素的一个快作用类似物，设计为模拟人体胰岛素对食物的反应的自然速率。改变 C 端的设计是基于与胰岛素样生长因子 1（IGF-1）结构和序列同源，序列颠倒减少了 B 亚基的二聚作用，使其更易于发挥作用，因此可以在进餐前很短的时间注射。自 1996 年起 Humalog 已经被许可使用。

二、结构域改组

大部分蛋白质由较小的独立的折叠域或结构域组成。这些结构域常通过短肽序列连接，并通常对应于其基因序列中的外显子。应用标准的分子生物技术，有可能对这些结构域进行改造，可通过添加（连接）、去除和交换结构域来重构蛋白质。

1. 连接域

（1）细胞靶向区域融合　在蛋白质工程非常早的例子中，一个抗体的结合位点区域通常与一种酶联系起来。抗体（免疫球蛋白）是一个 Y 形结构，由 2 条相同的 50 kDa“重”链和 2 条相同的 25 kDa“轻”链构成。抗体中的抗原结合区，被称为 Fab 区，由一条重链 N 端的一半和一条轻链组成。两条重链 C 端的一半联系在一起形成 Fc 区。后者与免疫系统的不同细胞间的相互作用及补体密切相关。至少对于 IgG 型免疫球蛋白，Fc 区在决定抗体血清半衰期方面起到非常重要的作用。

抗体-酶基因融合是这样进行的，将编码抗体 Fab 重链部分的 DNA 序列与期望连接的酶的编码序列连接后，导入产生抗体轻链的细胞中，并表达这些融合基因，使表达的融合蛋白具有抗原结合和酶活性的双重功能。类似的设计，也可利用抗体的 Fc 区与某种活性分子构建成融合蛋白，以利用抗体的长半衰期来改善所设计分子的药物动力学特性。最近一个已

经被许可对人类使用的例子是 Enbral™。它是由肿瘤坏死因子 α（TNF-α）受体的细胞外区域与 IgG 的 Fc 区融合得到，用来封闭 TNF-α 活性。

（2）细胞因子融合　不同细胞因子的功能经常有交叉，或者对相同细胞具有协同作用。通过构建融合蛋白，将两个功能活性单位连在一起，在理论上能够产生所期望的生物学效应。早在十多年前就将这种类型的融合应用于干扰素。最近报道了粒细胞巨噬细胞集落刺激因子（GM-CSF）与白介素 3（IL-3）的融合 PIXY32。在这个例子中，GM-CSF 和 IL-3 基因经修饰去除了哺乳动物 N 端糖基化位点，然后在 GM-CSF 的 C 末端和 IL-3 的 N 端之间通过附加一个 15 个氨基酸的柔性接头将 GM-CSF 与 IL-3 连接起来。与两个起始单体蛋白质的任何一个相比，酵母表达的 PIXY321 具有更强的受体亲和性、增殖活性和集落刺激活性。

（3）稳定二聚体蛋白的融合　在另外一些情况下，对在正常状态下是通过非共价作用或二硫键连接在一起的结构域可通过基因融合连接在一起。这使得生产更方便，并确保即使在低浓度时结构域保持在临近。抗体的结合域已经与接头肽连接在一起形成单链抗体（scFv）。与此相似，也能将细胞因子亚基连接到一起，IL-12 异二聚体已经被连接到一起并表明保持了功能。

2. 交换蛋白质结构域

蛋白质工程中另一个简单的过程是全结构域的交换，将来自不同蛋白质的类似结构域交换到一个多结构域蛋白质，从而提供了一个新的功能。

将来自鼠的单克隆抗体的抗原结合域连接到来自人的抗体的恒定区，构成嵌合抗体。与鼠源抗体相比较，这个转换显著减少不需要的免疫原性。到目前为止，已经有多个嵌合体抗体产品被许可作为药物，包括抗血小板复合物 ReoPro™，它是人-鼠嵌合体抗体的 Fab 片段。

3. 缺失结构域

组织型纤溶酶原激活剂（tPA）是一种由内皮细胞分泌的丝氨酸蛋白酶。它结合到血纤蛋白后，激活纤溶酶原成为纤溶酶，然后纤溶酶开始局部的血栓溶解。Reteplase 是一个 tPA 变异体形式，tPA 的五个结构域中有三个缺失，保留了与血纤蛋白选择性结合的结构域和催化结构域。Reteplase 可用于治疗急性心肌梗死，改善血液在心脏中的流动。

三、全蛋白的重组

许多蛋白质都是多成员家族的一分子，在家族中同源蛋白质的生物活性略有不同，这种不同是由序列的差异决定的。在蛋白质工程的早期，杂合基因是通过用合适的通用的限制性酶切交换序列段而产生的，或者是通过用更随机的方式在同源基因间进行体内重组而产生的。这些方法产生少量的新基因，并能够个别地被检测。最近更有用同源蛋白质通过序列重组来产生大的新变异体文库。在这个过程中，脱氧核糖核酸酶（DNase I）将家族的两个或更多个成员的基因随机分段，然后这些片段被作为 PCR 引物来扩增这些片段以便通过随机杂交产生新的基因序列，再通过前面介绍的筛选方法获得新的变异体。

四、蛋白质-配体相互作用

1. 酶修饰

蛋白质改造的许多例子都是针对酶的修饰，以检测或改变酶与底物的相互作用特性。这些实验可以追溯到最初的用酪氨酰-tRNA 合成酶进行的蛋白质改造尝试。

这类改变包括：增强催化活性；底物特异性修饰，包括新的催化作用的从头产生；改变 pH 谱使得酶能够在非生理条件下发挥作用；改善氧化抗性，方法是分别用空间相似的非氧化氨基酸如丝氨酸、苯丙氨酸或谷氨酸，取代氧化敏感的氨基酸如半胱氨酸、色氨酸或甲硫氨酸；通过取代 Cys 和 Met 残基和表面羧基来改善对重金属的稳定性；去除蛋白酶切割基元（motif）；去除催化产物可能结合而诱导别构反馈抑制的位点等。

最显著例子的就是工业生产用酶的使用范围和耐受性修饰。目前已经对大量的商业用酶进行了改进，包括蛋白酶、淀粉酶、脂肪酶、纤维素酶、木聚糖酶。其中一例，来自枯草蛋白酶 BPN′的不依赖钙的稳定性经修饰后提高了 1000 倍，改善了它在螯合环境中发挥作用的能力。

2．激素拮抗剂

提高激素与受体的亲和性能够导致发展具有显著增长的生物活性的超拮抗剂。近来，有人报道了通过合理的设计获得了活性增长 1300 倍的人促甲状腺素（TSH）变异体。该变异体是根据与人绒毛膜促性腺素（HCG）的同源性设计的，HCG 是共享一个公用的 α 亚基的另一个糖蛋白激素。替换正电荷基团到 TSH 的一环区使得活性协同增长。

3．结合特异性替换

有很多一个配体的特异性被从一个蛋白质背景转移到另一个蛋白质的例子。一些例子只需要小改变，一些则需要许多残基大规模的转移。例如，激素促乳素的受体结合表面的 8 个氨基酸被替换，特异性被修饰，所以现在这个被修饰的激素结合到生长激素的受体。

蛋白质结构域经常有环区，跨越其他的二级结构基元（α 螺旋，β 折叠）。在有些例子中，大蛋白质的功能区位于这些环区，并且这些环区也是相对简单的替换实验的靶位。例如，一个特殊环区的替换，使得碱性成纤维细胞生长因子的特异性被改变成酸性成纤维细胞生长因子特异性。

另外，抗体的人源化改造中也包括抗原结合结构域的转移，即将一个非人类抗体的抗原结合结构域的 β 折叠构架，转变成人的抗体的相应构架。自从 1986 年的首次描述，这项技术经过了许多年之后才成为一种常规。然而，现在已经发展到有大量的人源化抗体作为药物上市。

第四节　单克隆抗体的改造

抗体分子的若干独有的特性使其在疾病的诊断、治疗及预防等方面具有广泛的用途。抗体分子对抗原的亲和性高、特异性强，并具有特殊的生理效应，如补体激活作用与抗体介导的细胞毒作用等。然而，通过杂交瘤方法产生的鼠源单克隆抗体应用时有几个限制因素。首先，受 IgG 亚型的限制，鼠源单克隆抗体只能对人体免疫系统效应器功能产生弱的激活作用；其次，鼠源的抗体在人的血浆中有非常短的半寿期；最后，鼠源抗体容易引起人体的免疫反应，产生抗鼠源抗体的抗体来中和鼠源抗体的活性，这种人抗鼠抗体的产生可使人产生包括高过敏性在内的各种严重的副反应。

为了避免人抗鼠抗体反应，利用人源的单克隆抗体是首先考虑到的选择。然而，开发人

的杂交瘤工作并不成功。此外，用对人有害的癌细胞标记物作为抗原来对人进行免疫有悖伦理。因此，就促使人们应用蛋白质改造的方法产生嵌合抗体和人源化抗体来减少鼠单抗的免疫原性。经过改造，目前已经开发出多种适合不同用途的新型抗体，更适于临床应用。

一、抗体的产生和抗体的结构

1. 多肽构件的随机组合产生大量不同抗体

抗体是免疫系统的一个成分，抗体的主要作用是特异性地结合抗原并将其标记，让免疫系统将其消灭之。对于人体来说，大约有大于 10^8 个不同的抗体，如果每个抗体分子都有一个自身的基因那是太浪费了，因为这样所要的遗传信息量要超过人类基因组所含的信息量。包括人类在内的脊椎动物通过巧妙的方法解决了这个问题，这就是像盖房子一样，用少数几种标准的构件盖出千姿百态的建筑物来。各种抗体就是由标准的多肽构件连接而成的，在基因组中只有几百个这样的多肽构件的信息编码。少量大的构件为抗体的恒定结构域编码，这个恒定结构域决定了抗体的效应器功能，即保证了信息传递到免疫系统。然而，抗体的抗原结合特异性仅由抗体分子中的一小部分（可变区）来决定。这些可变区由两个到多个不同的模块组成，在 B 淋巴细胞分化的过程中，这些模块在每个细胞中以不同的方式结合到一起。除了这些基因片段的重组之外，也存在一些随机机制，即在各种基因片段的融合点随意地插入一些短的、新的序列。抗体各种可能结构的多样性也通过另外方式进一步增加，即抗体分子由两个独立的蛋白链（轻链和重链）结合成一个蛋白质复合体，而这两个蛋白链都参与抗原结合。图 9-8 给出一个免疫球蛋白 G 的结构示意图。

2. 抗原结合的特异性是由超可变结构域决定的

当将不同抗体分子的氨基酸序列进行比较时，发现在每个可变区内由 3～15 个氨基酸残基组成的三段肽序列具有非常大的可变性，故将这些区段称为超变区，它处于抗体分子的外表面，3 段来自于轻链，3 段来自于重链。超变区代表着抗体与抗原的实际接触位点。因

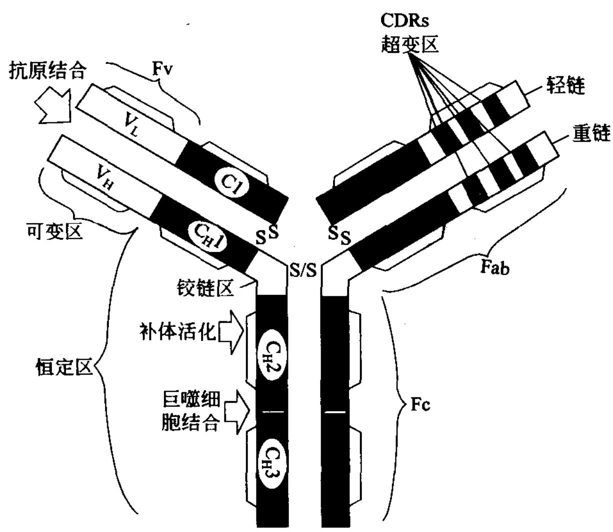


图 9-8 免疫球蛋白 G（IgG）分子及抗原结合片段 Fab、单链 Fv 的结构示意图

CDRs—互补性决定区；C1、C_H1、C_H2、C_H3—恒定区；V_L—轻链可变区；V_H—重链可变区；S/S—二硫键

为它们形成与抗原互补的结构，所以这些区域又称互补性决定区（CDRs）。而可变区其余部分的作用是稳定超变区的空间结构，它们是通过高度的刚性β片层结构形成的骨架来实现的。

3. 恒定区的作用是稳定抗体分子

抗体分子轻链和重链的两个可变结构域形成通常的抗原结合位点。这些结构域之间单独进行偶联通常相对较弱，因此要稳定抗体分子还需要其他的键。这些稳定抗体分子的键来自于恒定区。在恒定区的轻链、重链之间的二硫键可使其共价偶联；另外的二硫键将重链的恒定区相连成为一个二聚体，使抗体分子有两个一样的抗原结合位点。而在某类抗体中，多个这样的分子复合体被结合到一起，如 IgA 分子有 4～6 个抗原结合位点，而 IgM 则有 10 个抗原结合位点。

4. 恒定区介导效应器功能

只有在少数特殊情况下抗体靠自身行使其保护性功能。正如抗体可以防止病毒入侵目标细胞一样，中和性抗体可以抵抗细菌毒素。通过激活其他的免疫系统，共同参与免疫反应，是体液免疫最有效的作用方式。这种激活是由恒定区所介导的，而可变结构域则是结合和标记目的抗原。

不同的恒定区可介导一整套不同的生物学效应，如过敏反应是伴随 IgE 结合而产生的，而 IgM 结合可引起补体系统激活。通过模块设计可以交换恒定区而保留抗原特异性，根据需要，在免疫应答反应发生的过程中，免疫系统可以非常柔性的方式发生反应。

二、多抗、单抗及重组抗体

特定抗原在免疫实验动物中所产生的抗血清是含有不同结合特异性抗体蛋白的混合物。大体来讲，抗血清中绝大多数抗体并不与新的抗原结合，它们在用此抗原免疫前就已存在于血清中。新形成的抗体通常代表着不同抗原特异性的抗体混合物。每一种抗体对应于该抗原表面的一个抗原决定部位。因此，具有多种抗原决定部位的分子，在免疫动物后将产生多种抗体，即多克隆抗体（多抗），其中的每一种抗体都是由一类 B 淋巴细胞的克隆所产生的；针对一个抗原上单一抗原决定部位的特异性抗体，称为单克隆抗体（单抗）。

重组抗体则是通过基因工程手段在细菌或在细胞培养中取得。对这类抗体，关注的焦点是抗体的抗原决定部位。改造时通常将抗体分子的其他部位去掉，而将其抗原决定部位与酶蛋白或其他抗体等融合，这样得到的重组抗体就具有新的功能。与实验动物免疫和杂交瘤方法比较，重组抗体有着明显的优势：①由于可直接用工程菌生产而不是用脊椎动物，避免了 AIDS 和肝炎病毒等的污染；生产抗体的过程也比杂交瘤方法简便，便于培养、分析、修饰和贮存；②重组抗体是从单个细菌克隆来制备的，因此重组抗体是单克隆抗体；用大肠杆菌基因文库产生抗体不需要预先免疫，可以直接产生人抗体；③用基因工程技术将抗体分子与其他蛋白质融合，产生全新特性的抗体。例如，借助抗体片段可将一酶蛋白特异性地定位于肿瘤细胞上；通过设计可将两个具有不同特异性的抗原结合位点连在一起，形成一个双特异性的抗体分子，如果其中一个抗体片段能结合肿瘤细胞，而另一个能激活免疫系统，就形成一个可用于对肿瘤进行免疫治疗的双功能抗体。

现在，人们已经能用各种基因工程的方法组建抗体文库，表达抗体分子片段。为适用不同的目的，目前已经开发生产了多种改造过的抗体，如人-鼠嵌合抗体、小分子抗体及双功能抗体等。抗体改造成为蛋白质改造中最活跃的领域。

三、抗体改造

抗体改造的出发点是改善抗体的特异性、亲和性以及在受体细胞中的可生产性，当然也包括使其特性扩展、可同时作用于不同的抗原。对于很多应用而言，只改变亲和性是不够的。在一些应用中，特别小的抗体片段是必需的；鼠抗体必须改成人抗体；再者，增加抗体分子对蛋白酶的稳定性以及正确折叠都是重要的考虑。

（一）人-鼠嵌合抗体

如上所述，由于人抗鼠抗体免疫反应的存在，因此直接在治疗中应用具有特异性结合活性的鼠源单抗就受到限制。由于绝大多数人抗鼠抗体主要针对鼠源抗体的恒定区，使人们想到产生嵌合抗体可能解决上述问题，即将鼠单抗的恒定区换成人抗的恒定区。由于人抗体基因和鼠抗体基因都可以分别进行克隆，利用 PCR 技术很容易将鼠单抗可变区与人抗体的恒定区重组到一起。如图 9-9 所示，所得到的嵌合抗体仍然能特异性地结合抗原，而鼠源性降低。由于这样的嵌合抗体的恒定区（或恒定结构域）来源于人，因而它对于激活人免疫系统的某些辅助功能更有效，目前已经用到临床实践。

上述改造虽大大降低了鼠源单克隆抗体的免疫原性，但仍存在着少数抗鼠可变结构域的人抗鼠抗体。从抗体的结构图中可以看出抗体的可变结构域可以分成为超变区（即互补性决定区，CDRs）和变化性较少的骨架区，而与抗原的结合是由超变区决定的。与超变区相比，这些骨架区是保守的，这些有限的骨架区的氨基酸序列由一个紧密相关的基因家族编码。这种结构特点使人们想到，是否可以将鼠抗体的超变区嫁接到人抗体可变结构域的骨架部位，这就是所谓的超变区嫁接的设计（图 9-9），这样产生的抗体称为人源化嵌合抗体，也称为改形抗体或 CDR 移植抗体。

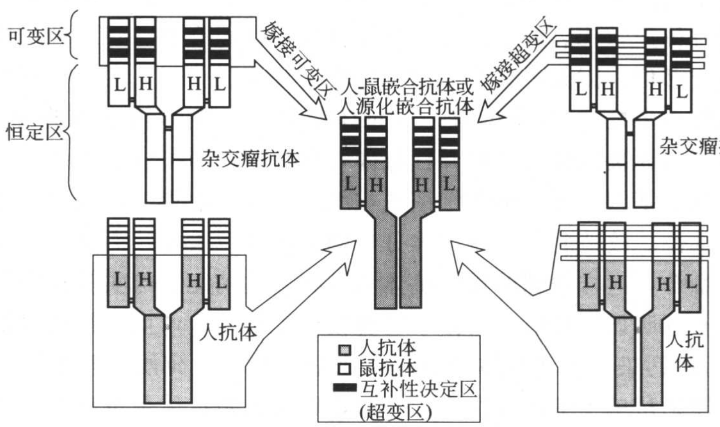


图 9-9 人-鼠嵌合抗体与人源化嵌合抗体的组建示意图

在构建改形抗体时，遇到的最大问题是抗体亲和力下降，甚至丧失活性。其原因是人 Ig 分子的框架区中的一些氨基酸与鼠 Ig 的 CDR 区不协调。Jones 等学者将鼠和人 Ig 的 V 区序列与其所属家族的保守序列进行比较分析后发现，如将鼠可能发生突变的残基、与亲和性有关残基引入人 Ig 的框架区中，可显著提高抗体的亲和力，成功构建出 CDw52 等多种改形抗体。或者是选择与鼠源性单克隆抗体同源性最大的人 Ig 分子为框架区序列，再将人 Ig 中较为独特的残基替换成鼠的 CDR 区附近的框架区的残基，抗体亲和力比原来鼠源性抗体

提高 30%，利用本方法成功地构建出 IL-2 受体的改形抗体等。随着对许多不同抗体的三级结构认识的扩大，CDR 嫁接所得到的结果也随之改善。现在，几个完全人源化抗体的亲和性已与鼠抗体几乎没有差别，用这样的抗体作为治疗制剂也再不受人抗鼠抗体反应的困扰。

（二）小分子抗体

基因工程小分子抗体仅表达鼠源性单克隆抗体的 V 区片段，其分子量仅为原抗体的 1/80～1/3，也就是说，保留了 CDR 区，而 Ig 分子的 C 区不表达。

现在研制的小分子抗体有以下几种类型：①Fab 片段抗体，由完整的轻链和重链（V_H＋C_H1）所组成；②FV 抗体，由 V_H 和 V_L 组成，天然 FV 片段中 V_H 和 V_L 为非共价性结合；③单链抗体（single chain FV，SCFV），由抗体 V_H 和 V_L 通过一段连接肽（linker）连接而成的重组蛋白。因为它能较好地保持抗体亲和力，具有分子量小、穿透力强、免疫原性小的特性，且易与效应分子相连接，构建出多种新功能的抗体分子，是构建免疫毒素和双特异抗体的理想元件，因此，SCFV 成为基因工程抗体研究领域中的热点；④单域抗体，它是由 V_H（或 V_L）单个可变区组成的，它只有抗体分子的 1/12，而且表面疏水性强，与抗原非特异性结合的能力也强；⑤最小识别单位（MRU）是由单个 CDR 区构成的小分子抗体，也具有与抗原结合的能力，但亲和力极低，实际应用有很大的局限性。

单链抗体的优点有：①可以去除许多造成非特异反应的竞争性表面蛋白，肿瘤显像的背景就会更加清晰；②更容易渗透到肿瘤组织中去，可以增加有效药物的治疗浓度；③分子量越小，鼠源性抗体的免疫原性越小，能消除人抗鼠的排异反应，延长在人体内的半寿期。单链抗体基因构建的一般方法是采用从杂交瘤细胞提取 mRNA，逆转录成 cDNA，通过 PCR 扩增抗体 V_H 和 V_L 基因，人工合成一条寡核苷酸序列（称 linker 序列），将 V_L 的 C 端与 V_H 的 N 端或 V_H 的 C 端与 V_L 的 N 端相连接，构建成单链抗体基因。近来，将噬菌体表面呈现技术应用于单链抗体筛选，可从外周血淋巴细胞直接提取 mRNA，绕过制备杂交瘤，甚至绕过免疫途径和筛选单链抗体基因。所使用的连接肽 linker 必须能使重、轻链可变区自由折叠，使抗原结合位点处于适当的构型，并应尽可能减少蛋白酶攻击和防止单链抗体的聚集。一般认为 linker 肽段长度以 14～25 个氨基酸残基为宜，目前最常用的序列是具有刚性结构的 15 肽序列 $(Gly4Ser)_3$。

单链抗体虽可较好保持抗原结合活性，但当抗原有重复抗原决定簇时，其结合活性会大幅度下降。通过构建双价 SCFV，可改善其亲和力。通过单链抗体的 C 末端连接另外单链抗体基因，构建和表达双特异单链抗体，可应用于肿瘤的诊断和治疗。它还可与 IL-2、TNF 等细胞因子基因连接，使细胞因子能特异性杀伤肿瘤细胞，增强抗肿瘤活性，降低毒副作用。

（三）双功能抗体

双功能抗体就是双特异性抗体（bispecific antibody，BsAb）。它是一种非天然性抗体，其结合抗原的两个臂具有不同的特异性。双功能抗体的构建有 3 种方法，即化学交联法、生物学方法（杂种-杂交瘤）和基因工程方法。

1. 化学交联法

1984 年 Karpovsky 等用琥珀酰亚胺吡啶二硫酚丙酸（SPDP）随机交联两个 Ig 分子表面上的巯基，构建出双功能抗体。1985 年 Brennan 等将 Fab 抗体绞链区巯基激活，构建出 Fab 双功能抗体。这种方法快速、简便、高效，而且不需要复杂的纯化过程，分子量有所减少，适合临床应用，因而得到了广泛使用。但此法构建的双功能抗体，因化学交联可能破坏抗体的功能，不太容易到达靶向部位，体内稳定性相对较差，且无连续性，每次应用都需要重新构建。

2. 生物学方法（杂种-杂交瘤）

1983 年 Milstein 等将分泌目的单克隆抗体的杂交瘤细胞与分泌另一种抗体的脾细胞融合，形成可分泌两种亲代抗体和杂种抗体的杂种-杂交瘤细胞，得到可预先设定的特异性的双功能抗体。该方法后来成为通行的构建双功能抗体的杂种-杂交瘤法。利用这种方法构建的双功能抗体具有严格的 Ig 的分子结构和相似的各种功能，并且可在一段时间内连续地提供双功能抗体；但构建过程复杂，周期长，产量低，纯化技术复杂，基因组过大且不稳定，又不能对双功能抗体进行改造，在一定程度上限制了它的临床应用。

3. 基因工程方法

1993 年 Holligen 等用基因工程方法构建出小分子双功能抗体（SCFV)$_2$，它是双功能抗体研究领域的里程碑。他们用接头连接不同抗体的 V_H 和 V_L 区，使它们不能形成链内的分子内配对，让其形成原抗体的分子内配对，构建成双特异性（SCFV)$_2$。(SCFV)$_2$ 是一种分子量较小的双功能抗体分子，不需体外组装，纯化步骤简单。最近又构建出双功能单链抗体，该方法把两种抗体的 V_H 和 V_L 基因连接在一起表达，形成的是一条融合蛋白链，既可避免同聚体的产生，又能大幅度提高产量。

利用基因工程方法构建的双功能抗体有很多优点。它既避免了化学交联对抗体的损害，又可定向结合抗原部位，减少了非特异性作用。1994 年又有学者提出了嵌合变构双功能抗体的概念，即把某种抗体的互补决定区 CDR 移植入 SCFV 抗体中，使其具有两种抗原的亲和性。虽然基因工程小分子双功能抗体尚未进入临床应用阶段，但其潜在的价值已引起重视，特别在肿瘤治疗中的应用前景格外引人注意。双功能抗体一个臂识别肿瘤细胞表面抗原，包括肿瘤相关抗原、癌基因产物、特殊的白细胞分化抗原、病毒蛋白抗原、特异性受体等，并实施有效的结合，而另一个臂可识别免疫效应细胞及分子，如 CD3、毒素、药物等，介导 T 淋巴细胞、LAK 细胞或毒素与药物发挥细胞毒作用。因此，相信双功能抗体在治疗肿瘤以及其他相关疾病（如自身免疫病、病毒性疾病等）的治疗中将发挥巨大的作用。

（四）抗体融合蛋白

抗体融合蛋白在广义上来讲应属于双功能抗体范畴，它是将抗体与另一种蛋白质连接构建成融合蛋白的形式。抗体融合蛋白的种类很多，如配基-配基型融合蛋白（如 CD4-Fcr）、配基-Ig 型嵌合蛋白（如 IL-2-Ig），或酶-抗体型融合蛋白（即抗体酶，abenzyme）。其中，较为成熟的是配基-Ig 型嵌合蛋白，而抗体酶则在药物治疗中应用前景广阔，如利用抗体介导的酶前体药物治疗，可通过抗体的特异性结合作用，将药物或药物前体定位到治疗部位从而提高了治疗区域的药物浓度，降低对正常细胞的毒性，使药物治疗更加精确。

构建抗体融合蛋白的基本方法并不复杂，可采用标准的融合蛋白质构建方法，但应当注意第一个蛋白质与第二个蛋白质之间应有一个适当的连接序列 linker，这是构建中的关键。一般来说，3～5 个氨基酸就可满足大部分融合蛋白的正确折叠的要求，如加入一段有疏水性和一定伸展性的较长肽键，可缓解两种蛋白的相互干扰。至于将哪一种蛋白质构建在 N 端或 C 端，通常没有明确的区别。目前基因工程抗体与 IL-2 的融合蛋白在肿瘤治疗中已取得较好的成绩，它可特异性将 IL-2 的作用集中在肿瘤局部，以减少 IL-2 对全身的副作用。Gillies 等构建抗神经节苷脂抗体/IL-2 的融合蛋白，在肿瘤局部的 IL-2 浓度比常规应用高 6 倍；构建的抗人表皮生长因子受体人鼠嵌合抗体与 IL-2 融合蛋白，可增强免疫效应细胞对黑色素瘤细胞的细胞毒作用。

（刘志国）

第十章 测序及人工合成技术

核酸和蛋白质是生物体内主要的生物大分子，各自有其结构特征，并分别行使不同的生理功能。核酸的功能是储存和传递遗传信息，指导和控制蛋白质的合成。遗传信息蕴藏在核酸分子的碱基序列中，测定核酸分子的一级结构，合成人类所需的基因或核酸片段是分子生物学研究的重要课题。

蛋白质是基因表达的最终产物，是细胞结构的组成成分，生命现象中绝大多数生物功能是通过蛋白质来实现的。测定蛋白质中氨基酸序列对于阐明蛋白质结构与功能的关系，阐明遗传病发病的分子机理以及采用基因工程定向改造生物具有重要意义。随着生命科学的发展，新的多肽如新的激素、释放因子、抑制剂、生长因子、生物毒素、离子载体和抗生素不断被发现，一些激素原和前激素原被分离。并且通过 DNA 测序，许多新的未知功能的多肽和蛋白质的序列在核酸的研究中被推测出来。这些研究迫切需要化学合成多肽和小的蛋白，来研究它们的构效关系和作用机理。此外，通过化学合成可获得大量天然难以获得的多肽和蛋白质，可获得具有免疫原性的亚单位疫苗。

第一节 DNA 序列测定

DNA 序列测定是分子生物学中的一项重要技术，是对 DNA 一级结构进行测定，即对 DNA 分子中的核苷酸顺序进行测定。由于它的建立和发展，使得人类基因组计划的实施成为可能。人类基因组计划的任务就是研究基因组的结构，即绘制遗传连锁图、物理图、序列图和转录图。其中，测序是最主要的任务之一。不仅如此，基因的分离、定位、转录、基因表达调控、基因工程载体的组建、基因片段的合成和 DNA 探针的制备等，都与 DNA 序列测定有密切的关系。

DNA 序列的测定技术是在高分辨率变性聚丙烯酰胺凝胶电泳技术的基础上建立起来的。变性聚丙烯酰胺凝胶电泳能分离长达 300～500 个核苷酸，而差别仅一个核苷酸的单链寡核苷酸片段。对于一个待测序列的 DNA 片段，需要使其转变成一系列的放射性核素标记的单链核苷酸，并使其一端为一固定的末端，而另一端由于长度不同，成为一系列相差一个碱基的连续末端。无论化学裂解测序法还是酶促合成测序法，在 4 种反应体系中，寡聚核苷酸分别终止于不同位置的 A、T、G、C 碱基。4 种反应体系的寡核苷酸产物，在变性聚丙烯酰胺电泳中上样于相邻的加样孔进行电泳分离。由于全部可能产生不同大小寡聚核苷酸都存在于 4 种反应产物中，从放射自显影的 4 种末端寡聚核苷酸梯子形图谱中，就可以直接读出 DNA 的序列（图 10-1）。

一、Sanger 双脱氧链终止法

（一）Sanger 双脱氧链终止法的原理

1977 年，Sanger 在加减法基础上创造了双脱氧核苷酸末端终止法。其原理是在酶促测

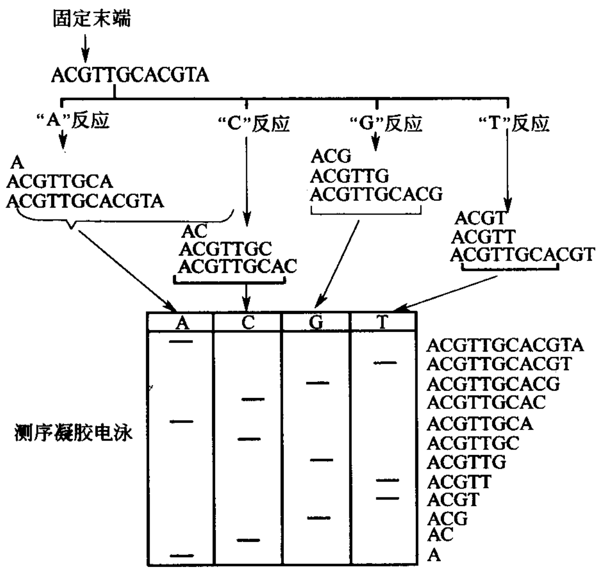


图 10-1 DNA 序列测定基本原理示意图

序反应中，利用一个与模板链特定序列互补的寡核苷酸作为引物来启动 DNA 的合成。在 DNA 酶促合成体系中，除了 DNA 聚合酶、DNA 模板、引物和 4 种脱氧核苷酸（dNTP）外，还加入一种双脱氧核苷酸（ddNTP），2′,3′-ddNTP 与普通 dNTP 不同之处在于它们在脱氧核糖的 3′位置没有羟基，它们可以在 DNA 聚合酶作用下通过其 5′-三磷酸基团结合到正在延长的 DNA 链中，但由于没有 3′-羟基，它们不能同后续的 dNTP 形成磷酸二酯键，因此，合成反应终止。

在 4 组独立的酶反应中，每个反应体系除了所加的 ddNTP 不同外，其他试剂均相同，即 DNA 模板、引物、DNA 聚合酶、4 种 dNTP，其中一种用 ^{32}P 标记或 ^{35}S 标记。如在 ddATP 终止的反应管中加 ddATP，在 ddCTP 终止的反应管中加 ddCTP，在 ddGTP 终止的反应管中加 ddGTP，在 ddTTP 终止的反应管中加 ddTTP。不过，这些 ddNTP 加入量都较少。当 DNA 合成时，这些 ddNTP 就与 dNTP 竞争掺入到新生的 DNA 链上，由于 ddNTP 比例较低，因此终止位点是随机的。经过适当的温育之后，结果将产生不同长度的 4 组寡核苷酸片段的混合物。无论是哪个反应体系，合成的 DNA 片段都具有相同的 5′末端，只是 3′末端不同。如是 T 管，合成的 DNA 片段由于掺入 ddTTP，而以 T 结尾；如是 A 管，合成的 DNA 片段 3′末端由于加入 ddATP，而以 A 结尾；同理，G 管，DNA 片段的 3′末端为 G；C 管，DNA 片段的 3′末端为 C。如果在反应体系中加入一种放射性标记的 dNTP，在反应后，生成互相独立的 4 组带放射性标记的寡核苷酸，然后在可以区分长度仅差一个核苷酸的高分辨率变性聚丙烯酰胺凝胶电泳条件下，对各组寡核苷酸进行电泳分析，只要把几组寡核苷酸加样于测序凝胶中若干相邻的泳道上，即可从凝胶的放射自显影片上直接读出 DNA 上的核苷酸顺序（图 10-2）。

（二）Sanger 双脱氧链终止法所需关键试剂

1. 模板与引物

两种类型的 DNA 可以用做测序的模板，即纯化的单链 DNA 和经过碱变性的双链 DNA。单链模板 DNA 常用噬菌体 M13mp 系列，使用单链模板能够得到比双链模板更好的结果。有经验的工作人员通过一次末端终止反应，能够读出 500 个以上的核苷酸序列。利用双链模板测序，若想得到比较好的结果，首先要制备高纯度的质粒 DNA，其次要选择高质量的酶。适合做双链模板的质粒，最好具有较高的拷贝数，有插入失活的选择标志，有配套供应的通用引物结合区，如 pGEM-3Z/4Z、pT7T3-18u/19u 等。以上两类质粒载体可分别从 Promega 公司及 Pharmacia 公司订购。

除特殊需要外，测序所用引物不必自己合成。根据所用载体的特点可从多家厂商购买到合适的引物。

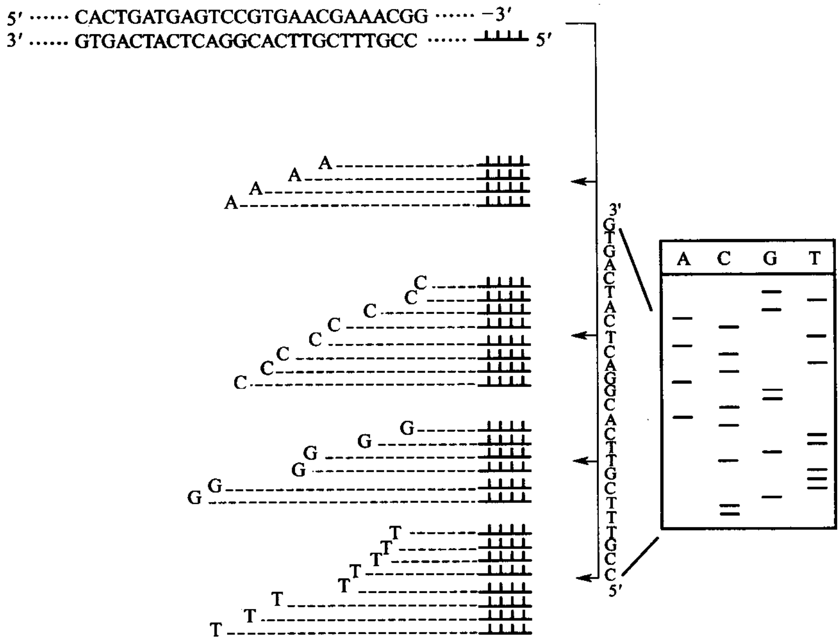


图 10-2　双脱氧测序原理示意图

2. DNA 聚合酶

测序所用的 DNA 聚合酶有 Klenow 酶、测序酶、*Taq* DNA 聚合酶等。

（1）大肠杆菌 DNA 聚合酶Ⅰ大片段（Klenow）　此酶是末端终止法最早采用的酶。用它进行测序常常产生较高的本底和假带，催化链延伸的反应的能力也比较差，通常只能读出距引物 250 个核苷酸以内的序列。此酶价格便宜，容易获得，对于已知序列 DNA 的亚克隆的鉴定仍有应用价值。

（2）测序酶　测序酶是经过改造的 T7 噬菌体 DNA 聚合酶，消除了 3′→5′外切酶活性。该酶活性非常稳定，具有很高的链延伸能力和极快的聚合反应速度，是测定较长 DNA 序列的首选酶。

（3）*Taq* DNA 聚合酶　该酶是聚合酶链反应（PCR）所需的关键酶，可在 95℃的高温下保持稳定。*Taq* DNA 聚合酶用于序列测定，除具有很好的链延伸性能外，还有可在高温（70～80℃）中进行反应的优点，因而能够克服富含 GC 序列模板形成自身二级结构对测序的影响。Promega 公司生产的“fmol™ DNA Sequencing System”试剂盒，将 PCR 反应与末端终止法测序结合进行，具有很多优点，例如只需很少量的模板，对双链模板的测序不需要碱变性的步骤，高退火温度增强了引物与模板结合的特异性，可以对多种模板如 PCR 产物、λ DNA、富含 GC 的模板及带有回文结构的模板进行序列测定。

3. 放射性核素标记的 dNTP

传统的测序方法都采用 α-^{32}P-dNTP，由于^{32}P 衰变过程中产生高能的 β 射线，导致放射自显影图谱条带扩散，分辨率低，限制了识别序列的数量。另外，由于^{32}P 对 DNA 样品辐射分解作用很强，通常都在测序反应后 24h 内上样电泳，否则无法得到满意的结果。

α-^{35}S-dATP 近年得到广泛使用，它解决了^{32}P 的两大缺点，具有很高的分辨率、较低的本底。测序反应产物－20℃保存 1 周并不明显影响结果。

（三）用于测序的变性凝胶电泳

DNA 序列测定的准确性很大程度取决于测序产物在变性聚丙烯酰胺凝胶上的分离度。一般说来，用于 DNA 序列测定的凝胶需要长 40cm，厚度均匀，含有 4%～8%丙烯酰胺和 7mol/L 尿素。一般推荐使用 6%丙烯酰胺，非梯度胶为起点。可使用两种胶，一种用于分离较短的寡核苷酸，另一种用于分离较长的寡核苷酸；如用 6%的胶分离较长的寡核苷酸，用 8%的胶分离较短的核苷酸。

为保证电泳过程的均匀散热，可将一片铝平板（0.4cm 厚，34cm×22cm）夹到凝胶夹层前面的玻璃平板上，但注意不要让它接触到电泳缓冲液。

（四）大片段 DNA 的测序策略

实际上，从一套测序反应中所能获得的信息量受限于测序胶的分离度。虽然最新的测序技术经常可从一套测序反应中得到高达 500 个核苷酸的信息，但获得的可靠信息大约在 300 个核苷酸。因此，如果待测 DNA 的区段在 300 个核苷酸以内，所需的工作只是简单地将此片段克隆于合适的载体，以产生一个能方便地进行测序的重组 DNA 分子。

对于大片段 DNA 的序列测定，往往需要将其切割成能单独进行测定的小片段，这可通过随机的或有序的方式分段进行测序，然后再汇合成完整的长链 DNA 序列。常用的方法有以下三种。

1. 随机法

又称鸟枪法。将长链 DNA 随机断裂成适于测序的片段，其手段有三种，即超声波处理、DNase Ⅰ（核酸酶）切割、限制酶消化。

（1）超声波处理　用超声波将 DNA 随机断裂成 1kb 左右的片段作亚克隆，成为一组含有互相重叠的克隆。

（2）DNase Ⅰ切割　在 Mg^{2+}存在下，DNase Ⅰ在双链 DNA 分子上随机形成缺口。但在 Mn^{2+}存在下，它可随机切割 DNA 的两条链，产生一组相互重叠的片段，将这些片段与适当的载体相连，即形成亚克隆。

（3）限制酶消化　用限制酶切割 DNA，可产生一组含黏性末端的 DNA 片段，将其与适当载体连接后，形成亚克隆。

将上述含有不同子片段的亚克隆扩增后分别进行 DNA 序列测定，再将序列重叠排列，便可形成大片段 DNA 的完整序列。无论是超声波处理，还是 DNase Ⅰ切割或是限制酶消化，产生的 DNA 片段的末端都难免有破坏。这样，测序时，有的子片段的序列出现多次，而有的子片段则始终不出现，致使出现大片段 DNA 序列测定不完整的现象（造成有时得不到大片段 DNA 的完整序列），而且测定的核苷酸数目比实际 DNA 全长多几倍。

2. 嵌套缺失法

又称互套缺失法（nested deletion）。该法是将大片段的待测 DNA 克隆入载体，然后用限制酶将其切成线性 DNA，用核酸酶 BAL3Ⅰ和核酸外切酶Ⅲ将线性 DNA 的一端删切 200～300bp（或更长），形成一系列删切 200～300bp 的 DNA 片段。再将这些大小不同的 DNA 片段重新环化，就构成了一系列一端删切 200～300bp 的 DNA 片段的亚克隆（图 10-3），再分别对它们测序，就可以连续读出该大片段的全部序列。其步骤如下。

① 将待测 DNA 与 M13mp 载体相连，构成重组体。

② 用两种限制酶从待测 DNA 片段的一端与载体序列之间将 DNA 切断，产生线性 DNA 分子。线性 DNA 分子近待测 DNA 片段的末端为 5′突出的黏性末端，而另一端则是近

载体的一端，为 3′突出的黏性末端。

③ 用外切核酸酶Ⅲ消化上述线性 DNA。外切核酸酶Ⅲ能从 DNA 的平端或 5′突出末端起始消化，而有 3′突出的双链 DNA 完全不被其降解。根据这一特性，设计出了单一方向删切线性双链 DNA 的方法。使其删切只发生在近待测 DNA 一端，把待测 DNA 的一端变为 5′突出的黏性末端，而另一端（载体）则变为 3′突出的黏性末端。这样核酸外切酶的删切就只发生在待测 DNA 一端，使待测 DNA 逐渐缩短。

核酸外切酶Ⅲ在 37℃时，每分钟水解 250 个核苷酸，速度较均匀。随着消化时间的延长，待测 DNA 较匀速地变短，采用不同时间终止消化反应，可以人为地获得一组 DNA 片段，这些片段在同一端缺失，依次相差 200～250bp，而载体（包括引物结合区）一端完全保留。再经过核酸酶 SⅠ 消化掉残余的单链，使两端均成为平末端。然后自身连接环化，就可直接转化或转染宿主菌得到亚克隆。

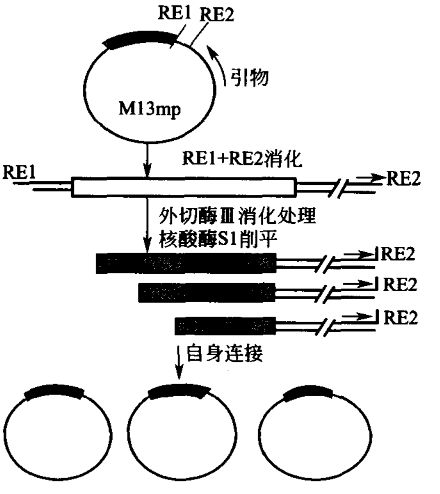


图 10-3　嵌套缺失法建立连续亚克隆测序

④ 将各个克隆用于测序。从各个片段的缺失端开始测序，测序中所用的引物相同，测序结果从子片段序列相互重叠的部分便可准确无误地将相邻片段的序列彼此拼接，如此重叠嵌套，便可推测出大片段 DNA 全长序列。

3. 引物延伸法

是一种定向测序法。从 DNA 的 3′端，依赖特定引物延伸可测定一段 DNA 序列，再根据测得序列设计新的寡聚核苷酸，作下一延伸反应的引物，如此逐渐向前推进，最终测得 DNA 的全长序列。该法优点是不必亚克隆待测 DNA 的子片段，整个测序过程以逐步推进方式进行。缺点是需要合成许多引物，费时、价格昂贵。

二、Maxam-Gilbert 化学修饰法

（一）化学裂解法测序的原理

1977 年，Maxam 和 Gilbert 建立了一种以化学修饰为基础的 DNA 序列分析法，称为 Maxam-Gilbert 化学修饰法或称化学裂解法。其基本原理：首先对待测 DNA 进行单侧末端的放射性核素标记。标记后的 DNA 分成 4～5 个反应体系，分别用不同的化学试剂处理，使 DNA 片段分别于某一种或某一类碱基处断裂，通过控制化学反应进行的程度，使得平均每个 DNA 分子只在一个位置被裂解，从而获得一系列大小不同的 DNA 片段，将这些片段经聚丙烯酰胺凝胶电泳分离，经放射自显影后得到互相错落的梯形图谱，即可读出 DNA 序列（图 10-4）。

例如在 G 反应体系中，用硫酸二甲酯使鸟嘌呤第 7 位 N7 甲基化，再用吡啶处理，DNA 在任意位置的鸟嘌呤处断裂，产生一组一端为放射性标记的固定末端，另一端为鸟嘌呤的不同大小的 DNA 片段混合物。它们的长度可以从数个核苷酸到接近待测 DNA 的全长。连同其他反应体系 A+G、C+T、C 的产物在变性聚丙烯酰胺凝胶上相邻的加样孔中进行电泳分离，放射自显影后得到互相错落的梯形图谱即可读出 DNA 序列。

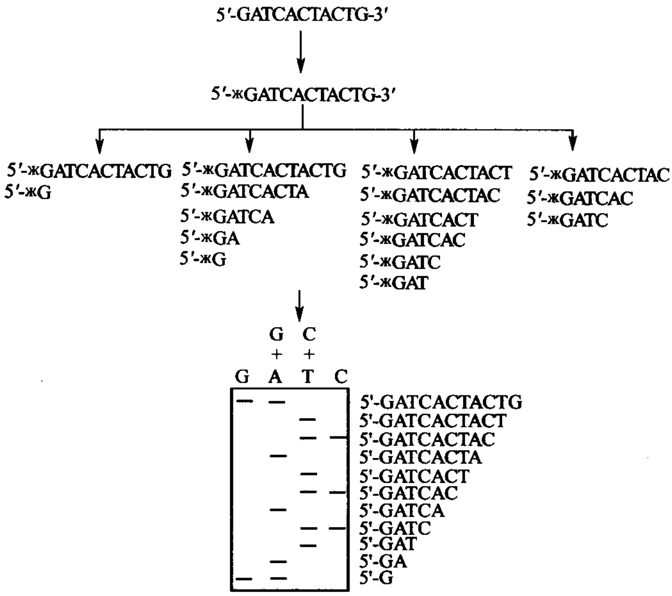


图 10-4　化学裂解法测序原理示意图

（二）化学试剂特异断裂 DNA 的机制

在 4 种反应体系中，化学试剂特异地断裂 DNA 机制如下。

G 反应：硫酸二甲酯是一种碱性化学试剂，既可使腺嘌呤 N3 甲基化，也可使鸟嘌呤 N7 甲基化，但两者甲基化修饰的速度不同，使鸟嘌呤 N7 甲基化的速度比使腺嘌呤 N3 甲基化的速度快 4～10 倍。在中性环境中，主要作用于 G；该碱基的甲基化，在加热的条件下，吡啶使其被修饰碱基处的磷酸二酯键断裂。

A+G 反应：甲酸使嘌呤环上的氮质子化，导致糖苷键被削弱，在吡啶存在时，嘌呤环被吡啶取代，吡啶在加热条件下导致 DNA 链在脱嘌呤位点（A 或 G）发生断裂。

G+T 反应：肼能够裂解嘧啶环，进而导致其脱落，在吡啶存在时，导致被修饰碱基处的 DNA 断裂。

C 反应：在一定浓度的 NaCl 条件下，只有 C 才能与肼发生反应，随后被修饰的胞嘧啶被吡啶置换。

在所有 4 个反应中，吡啶在加热的条件下，能催化被修饰碱基所在位置的磷酸二酯键断裂。

在化学修饰反应过程中，通过控制反应温度和反应时间，只有一小部分碱基被修饰（而不是全部被修饰），随后进行的断裂反应也是定量反应。因此，DNA 链并不是在所有被修饰的碱基位点断裂，而是随机断裂。在 4 个反应中，产生 4 套带相同标记末端长短不一的寡聚核苷酸片段。只有带标记末端的片段可被识别。没有标记末端的片段可以忽略不计。

4 个反应体系的混合物经高分辨率聚丙烯酰胺凝胶电泳及放射自显影，可直接读出待测 DNA 的碱基排列顺序。

三、DNA 序列的自动测序

DNA 序列测定是在 20 世纪 70 年代中期提出并逐渐完善起来的，Sanger、Maxam 和

Gilbert 等人分享了 1979 年的诺贝尔化学奖。自 20 世纪 80 年代以来，随着计算机软硬件技术，仪器制造和分子生物学研究的迅速发展，使得 DNA 自动测序技术取得了突破性进展。早在 1987 年 Perkin Elmer（PE）Applied Biosystems 公司就推出 DNA 自动测序仪，其专利是分别采用 4 种荧光染料进行标记而在同一泳道测序，具有极大的优越性。377 型全自动 DNA 测序仪是 PE 公司近年推出的新型测序仪，它采用专利的 4 种荧光染料标记，并采用激光检测方法，具有测序精确度高、每个样品判读序列长（700bp）、一次电泳可测定样品数量多（64 个）、不需要同位素测序、方法灵活多样等特点。在人类基因组测序和 cDNA 文库测序研究中应用极其广泛。此外，该仪器在各种应用软件的辅助下，还可以进行 DNA 片段大小分析和定量分析，应用于基因突变分析、DNA 指纹图谱分析、基因连锁图谱分析以及基因表达水平的研究，有着极其广泛的应用前景。

全自动 DNA 测序原理与 Sanger 双脱氧末端终止法相同。所不同的有两点：一是标记物不同，手工测序采用放射性核素标记，自动测序采用的是 4 种荧光染料，这 4 种荧光染料分别标记不同的终止物 ddNTP 或标记引物，将这些标记物分别加到不同的反应管中，经 Sanger 测序反应后，产物 3′端（标记终止物 ddNTP）或 5′端（标记引物法）带有不同荧光标记，一个样品的 4 个测序可以在一个泳道内电泳，从而降低了测序泳道间迁移率差异对精确性的影响；二是检测手段不同，手工测序采用放射自显影，从 4 种末端寡聚核苷酸的梯子形图谱中，直接读出 DNA 序列，而自动测序则是经电泳将各个荧光谱带分开，同时激光检测器同步扫描，一次扫描可检测出多种荧光，不同荧光代表不同的碱基信息，传入电脑，其测序速度高达 200bp/h。自动测序仪连续采集的数据在计算机中储存并迅速加以处理，结果清晰准确，分辨率很高，操作人员可以根据显示屏上的彩色显示图像准确分辨出某一 DNA 样品的碱基排列顺序。读序时，根据计算机软件中碱基的代表符号，进行阅读和编辑。

第二节　DNA 的化学合成

DNA 的化学合成已在分子生物学中得到广泛应用。早在 1958 年，英国剑桥大学 Todd 实验室首先合成了具有 3′→5′磷酸二酯键结构的 TpT 和 pTpT。此后，美国麻省理工学院 Khorana 等对基因的人工合成作出了划时代的贡献，他们在 1970 年，首次报道了酵母丙氨酸转移核糖核酸基因的人工合成，1976 年又成功合成了大肠杆菌酪氨酸转移核糖核酸基因。他们不仅创建了基因合成的磷酸二酯法，而且发展了一系列有关核苷酸的糖上羟基、碱基的氨基和磷酸基的保护基及缩合剂和合成产物的分离、纯化方法。随着 DNA 合成技术的发展，到目前为止，DNA 的合成方法有液相磷酸二酯法、磷酸三酯法、亚磷酸三酯法，以及在后两者基础上发展起来的固相的磷酸三酯法和亚磷酸三酯法及自动合成法。

从原则上讲，任何大小的基因和核酸片段都可用化学方法合成。但如果基因比较大，用化学法合成则十分费事。因此，对大分子量的基因而言，直接从生物体中分离天然的基因则比较适宜。人工合成的基因可以是生物体内已经存在的，也可以是按照人们的愿望和特殊需要重新设计的。因此，它为人类操纵遗传信息、校正遗传疾病、创造新的优良物种提供了强有力的手段，是基因研究的一个富有成效的重大飞跃。

一、DNA 化学合成的原理

DNA 的合成方法有液相合成法和固相合成法。液相合成法相对来说太费力，耗时且难掌握，而固相合成则具有简便、快速、通用和易操作等特点，特别是磷酸三酯法和亚磷酸三酯法与固相合成技术相结合，使 DNA 的化学合成面目一新，无论是手工合成还是自动合成仪合成都既省时又易于掌握。事实上，固相合成法已基本取代了液相合成法。下面重点介绍固相磷酸三酯法和亚磷酸三酯法合成的原理。

（一）磷酸三酯法合成寡核苷酸的原理

磷酸三酯法与磷酸二酯法一样，包括单体的制备、缩合和分离纯化等。它们都使用了在单核苷酸的 5′末端或 3′末端加保护基团的做法，以防止在这两个末端之间形成磷酸二酯键。但在磷酸三酯法中，参加缩合反应的单核苷酸，是一种核苷 3′-单磷酸的衍生物。在它的磷酸分子上有一个羟基（P-OH）已经带上了适当的保护基团（通常是邻氯苯基和对氯苯基），因此事实上已经是一种磷酸二酯，而余下的另一个 P-OH 仍然具有反应活性，可以同另一个带有 3′末端保护基团之单核苷酸的 5′-OH 缩合形成具磷酸三酯键的二核苷酸分子。化学合成的生长激素释放抑制因子基因（又叫脑激素基因），就是磷酸三酯法实际运用的一个成功例子。这个合成基因能够在大肠杆菌细胞中表达。但早期的磷酸三酯法合成基因的程序十分复杂，而且全部反应都是在液体中进行，因此难以控制。为了克服这方面的缺点，以后又发展出了在固体载体上进行寡核苷酸化学合成的方法，即固相合成法。

在固相合成中，第一个脱氧核苷酸是通过它的 3′-OH 基团直接附着在惰性的固相载体上。这种载体同样也起到一种 3′末端保护装置的作用。在实际合成过程中，固相载体（如硅胶）是被装填在一种烧结玻璃滤柱内。试剂只是在规定时间内才允许同末端脱氧核苷酸进行连接反应，并且每次只接长一个核苷酸或一个寡核苷酸短片段。在进行下一次的连接之前，要用适当的溶剂将载体清洗干净。这样，每连接一次就经历一次循环，合成的寡核苷酸链则始终被固定在固相载体上，而过量的未反应的试剂或分解的物质，则被过滤或洗涤除去，当整个链增长到所需的长度之后，再将所合成的寡核苷酸链从固相载体上切除下来，并洗脱保护基，经纯化得到所需要的终产物（图 10-5）。

（二）固相亚磷酸三酯法合成寡核苷酸的原理

固相亚磷酸三酯法，是目前最通用的一种合成寡核苷酸的方法。合成的原理及步骤见图 10-6。首先，将欲合成的寡核苷酸链 3′末端核苷 1（N1）以其 3′-OH 通过 1 个长的烷基臂附着在惰性的固相载体上，这种固相载体典型的情况是使用可控微孔玻璃（CPG）。同时，它的脱氧核糖环的 5′-OH 基团也已用二甲氧基三苯甲基（DMT）保护起来。然后，从 N1 开始逐步地接长寡核苷酸链。合成的一个循环周期分为如下 4 个步骤。

第一步为去保护（deprotection）。用酸处理法，脱去与核苷酸 1 的 5′-OH 基团偶联的保护基团 DMT。在这种脱 DMT 反应中，常用的酸是二氯乙酸（DCA）或三氯乙酸（TCA）。

第二步为偶联反应（coupling）。通过一种弱碱——四唑（tetrazole）的催化反应，使加进来的第二个核苷酸（N2）同已附着在固相载体上的核苷酸 1 之暴露的 5′-OH 基缩合。

第三步为封端反应（capping）。由加入的乙酸酐（acetic anhydride）激发的乙酰化作用，把所有的没有参与偶联反应的 5′-OH 基团全都封闭起来。

第四步为氧化作用（oxidation）。在核苷酸 N1 和 N2 之间新形成的 3′,5′-亚磷酸三酯键十分活跃。利用碘液催化作用，使之被氧化成为相当稳定的 3′,5′-磷酸三酯键。这也就是为

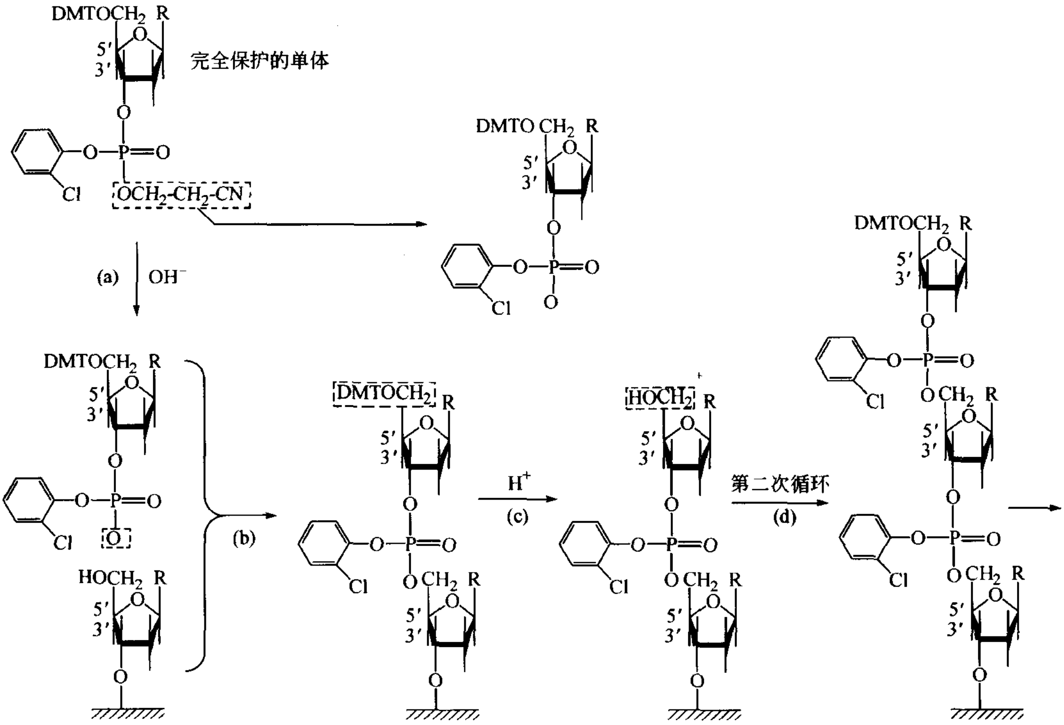


图 10-5　固相磷酸三酯法合成寡聚脱氧核苷酸片段

(a) 完全保护的脱氧单核苷酸衍生物，通过三乙胺（triethylamine）的作用，使其 3′-P 位置处于未保护状态；(b) 于是，它便可同 3′-OH 端被固定在固相支持物上的另一个脱氧单核苷酸分子上未保护的 5′-OH 缩合，形成二核苷酸；(c) 加酸处理，使其脱去 5′-OH 上的保护基团 DMT；(d) 这样，它又可同具 3′-活性末端的另一个脱氧单核苷酸分子缩合，形成三核苷酸分子。如此重复进行，到反应最后完成时，所有的保护基团都要被清除掉，并通过高效液相色谱法（HPLC）纯化合成产物

图中 R 分别为苯甲酰腺嘌呤（benzoyl adenine）、苯甲酰胞嘧啶（benzoyl cytosine）、异丁基鸟嘌呤（isobutyryl guanine）或胸腺嘧啶

什么称这种寡核苷酸合成法为亚磷酸三酯法的原因。

在下一个合成循环之前，要把附着在最后一个偶联核苷酸上的二甲氧基三苯基（DMT）移去，以保证它的 5′-OH 基团能够暴露出来，同一个活化的脱氧单核苷酸溶液进行偶联反应。DMT 是一种显色指示剂，所以它的释放可用来检测偶联反应的效率。

上述这种循环反应，可被重复地进行，直到合成出具有所需序列长度的寡核苷酸片段为止。到合成终止时，固相载体上携带着被完全保护的寡聚脱氧核苷酸，因此需要使它们有系统地去掉保护基团，并从固相载体上释放出来，成为游离的寡核苷酸。然后，用乙醇沉淀法纯化这些寡核苷酸，再用高效液相色谱法或凝胶电泳法使之进一步纯化，并同未完成的较短的片段分开，最后得到真正需要的产物。

固相亚磷酸三酯法的主要优点是合成产率高、反应速度快，如果使用更加活跃形式的磷，那么大约在 2min 内就可完成一次缩合作用。目前，此法已成功地用来合成长达 150bp 以上的寡核苷酸片段。因此，一般认为固相亚磷酸三酯法是目前比较理想的一种寡核苷酸合成法。

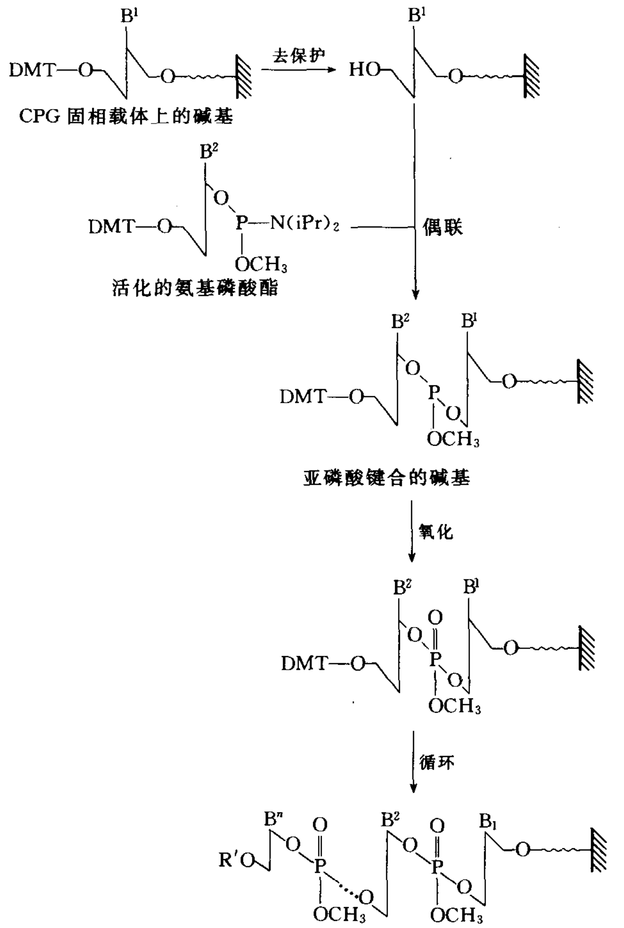


图 10-6 固相亚磷酸三酯法合成寡核苷酸片段

DMT＝二甲氧基三苯甲基（4,4′-dimethoxytrityl）

（三）用寡核苷酸片段组装基因的方式

目前，化学合成寡核苷酸片段的能力一般局限于 150～200bp 以内。然而，绝大多数基因的大小都超过了这个范围。因此，需要一种特殊的程序，才能够把合成的寡核苷酸片段构建成完整的基因。我们将这种按设计要求用许多寡核苷酸片段装配成完整基因的过程，叫做基因的组装。Khorana 和他的合作者，在 20 世纪 70 年代末最早提出基因组装法，并成功地合成出了 tRNA 基因。常用的组装基因的方法有两种。

第一种方法是先将寡核苷酸激活，带上必要的 5′-磷酸基团，然后再与相应的互补的寡核苷酸片段退火，形成带有黏性末端的双链寡核苷酸片段。把这些双链寡核苷酸片段混合在一个试管中，加上 T4 DNA 连接酶，使它们彼此连接组装成一个完整的基因或基因的一个片段。这些组装的产物被插入到适当的噬菌体载体或质粒载体上，并转化到大肠杆菌寄主细胞中。最后，用 DNA 序列分析法检测所组装的基因。应用这种基因组装法，已经构建了许多基因，其中包括分子量较大的 α-干扰素基因和牛视紫红质基因。

第二种方法为酶促合成法。首先，合成末端之间有 10～14 个互补碱基的寡核苷酸片段，退火后以重叠区作为引物，在 4 种 dNTP 存在的条件下，通过 Klenow 酶的聚合作用，便会迅速地合成出相应的互补链。如此形成的双链 DNA 片段，或是先用平末端直接同选择好的

载体分子连接，或是用适当的限制酶处理，使之形成黏性末端之后再插入到载体分子上。从理论上讲，应用这种方法组装基因，在一开始就加入具有重叠末端的寡核苷酸片段，能够合成出大分子量的基因，然而实际上由于在退火过程中难以消除那些不必要的杂交作用，同时二级结构的存在也会干扰聚合酶（Klenow 酶）的作用，这种基因组装法还是有相当的局限性。

利用长的寡聚体能够直接组装长度为 1500～2000bp 的基因。遗憾的是，用这种办法构建的基因，其突变频率相当高，平均每隔 500～800bp 就会出现一个突变。

二、DNA 合成技术在分子生物学研究中的应用

1. 合成基因

目前有许多基因/蛋白质的核苷酸/氨基酸序列已得到阐明，人们已经可以根据需要合成出具有实际应用和研究价值的多肽和蛋白质的基因。已报道的合成基因有酵母丙氨酸转移核糖核酸、人生长激素、干扰素、胰岛素、表皮生长因子、白细胞介素Ⅱ、集落刺激因子等基因，这些基因均已被克隆，绝大多数已在原核或真核系统中获得表达。

2. 合成探针

基因的克隆和分离已成为现代分子生物学研究必不可少的组成部分。DNA 合成技术在其中起着越来越重要的作用，它不仅使得过去颇费周折的基因筛选与鉴定成为常规技术，而且使得克隆载体的构建及克隆基因的载体的连接变得更容易和准确。蛋白质的一级结构可以通过 mRNA 的碱基序列间接地得出。相反，如果已知某个肽段的氨基酸顺序，也可根据密码简并的原则推导出所有可能的 mRNA 序列密码。人工合成的具有特定顺序的寡聚 DNA 片段，已应用于筛选和鉴定重组质粒或 λ 噬菌体。实验证明用合成的寡核苷酸片段，即使只有 1 对碱基错配，采用严谨条件杂交也能与完全互补的双链相区别。因此，作为探针使用的寡核苷酸，将密码的简并度调至最高限度时，可减少假阳性，增加筛选的准确率。

3. 合成引物

(1) 合成 PCR 引物　PCR 技术是近年发展起来的一种体外扩增特异 DNA 片段的技术，此法操作简便，可在短时间内在试管中获得数百万特异 DNA 序列的拷贝。其基本原理与细胞内 DNA 复制相似，但反应体系相对简单。它是在模板、引物和 4 种脱氧核糖核苷酸存在的条件下依赖 DNA 聚合酶的酶促合成反应。它是用于扩增位于两段已知序列之间的 DNA 区段的一种方法。PCR 技术的特异性取决于所用引物和模板 DNA 结合的特异性。合成引物在 PCR 反应中被普遍使用。

(2) 合成测序引物　DNA 序列测定是分子生物学中最重要的研究技术，其中最常使用的是末端终止法，它是一种依赖于特异 DNA 引物的序列测定方法，此外该法对模板的需要量较大，这就要求将待测 DNA 片段尤其是拷贝数少的片段首先克隆到适合的载体中经过扩增后进行序列测定。常用的克隆载体如 M13、pUC19、pBR322 等均可购到有商品出售的通用测序引物。

(3) 合成定点诱变的引物　无论是寡核苷酸介导的基因定点诱变还是 PCR 介导的基因定点诱变都是通过设计引物来完成。利用寡核苷酸引导的突变，可在目的 DNA 序列的任何部分产生点突变、插入和缺失，从而使得基因编码的蛋白质在结构和功能上发生改变。

4. 合成人工接头

在 DNA 重组中常需要将外源 DNA 片段插入某些载体 DNA 中。如果载体或外源 DNA

上没有合适的限制性内切酶位点，为提高插入效率或实现定向克隆，有效的办法就是采用合成人工接头的方法。使用人工接头需要插入片段是平末端，因此对于非平头的插入片段，首先需要用 DNA 聚合酶 Ⅰ 大片段补齐或者用核酸酶 S1 或 BAL31 处理，得到平末端后才可与人工接头连接。

具有多酶切点的接头（polylinker）还广泛应用于构建某些运载体，使之带有多个新的酶切位点。这在质粒 pUC 系列及用于序列分析的 M13mp 噬菌体系统中都得到了广泛的应用。接头片段还用来调整表达载体的读码框架，用于基因表达及功能的研究。

5. 合成基因表达调控序列

基因表达调控序列包括转录启动子、操纵基因、核糖体结合位点、增强子及转录终止序列等。利用化学合成方法合成这些序列有助于研究基因结构与功能的关系。

（1）合成启动子　启动子是位于基因的表达调控区的一段短的 DNA 序列。例如，原核基因的转录启动序列通常在 40 个核苷酸以下，核糖体结合位点在 20 个核苷酸以内，而转录终止序列小于 30 个核苷酸。这些都在有效的 DNA 合成序列长度以内。通过对大肠杆菌多个启动子序列的研究发现，在转录起始位点上游有 2 个保守区域，即 −10 区的 TATAAT 和 −35 区的 TTGACA，它们分别是 RNA 聚合酶的结合位点和识别位点，它们和转录的效应密切相关。采用 DNA 合成的方法，人们将 *lac* 启动子的 −10 顺序和 *Trp* 启动子的 −35 顺序融合组成杂合的 *tac* 启动子，它的转录效率比 *Trp* 启动子提高 3～5 倍。还可用人工合成的方法调整基因起始密码（ATG）和 SD 的序列之间序列的长度，可以获得高效的表达。

（2）合成操纵基因　操纵基因是原核操纵子上与阻遏蛋白结合的 DNA 序列。乳糖操纵基因是最先进行合成研究的调控基因。在乳糖阻遏蛋白和乳糖操纵基因相互作用的研究中，合成手段弥补了遗传分析和化学修饰技术的不足，应用核苷类似物在某一位置取代原有核苷的方法进行体外阻遏蛋白-操纵基因结合实验，帮助人们认识了 DNA 分子上功能基因的重要作用。在乳糖操纵基因的研究中，用 DNA 合成手段确定了其中 17 个核苷酸长的片段是其显示生物学活性的最短序列。

（3）合成核糖体结合位点　核糖体结合位点是原核 mRNA 能否作为翻译模板的决定因素。此调控区一般由含嘌呤的顺序组成，其下游为翻译起始密码，二者间有 4～10 个核苷酸的 SD 序列。以合成 DNA 片段取代的方法发现核糖体结合位点 GGTTAA 的存在，具有普遍的提高表达效率的作用，并且核糖体结合位点和起始密码之间的核苷酸数目和组成都能影响翻译的效率。

6. 合成基因在医学中的应用

合成的 DNA 片段已直接用于某些遗传病、传染病的诊断和基因治疗。

（1）用于遗传疾病的诊断　过去仅有少数遗传性疾病可以凭借蛋白质和酶学方面的异常作出判断，随着分子生物学的发展，对遗传病的诊断有了巨大的进步，出现了基因诊断法。例如，用合成的基因特异性寡核苷酸探针诊断镰刀状红细胞贫血、β-地中海贫血、α-地中海贫血、α1-胰蛋白酶（AAT）缺陷病、苯丙酮尿症等遗传性疾病。

（2）用于传染病的诊断　临床上对传染性疾病的传统诊断程序常常是检测病原体及其抗体，需时较长，达不到早期诊断的目的。合成探针在传染病特别是由病毒感染所致的传染病检测中，已得到广泛应用，尤其是结合 PCR 方法使诊断的灵敏度可进一步提高，目前，国内已研制出检测乙肝病毒、艾滋病病毒（HIV）、轮状病毒和疱疹病毒等的基因探针诊断试剂盒。

（3）用于基因治疗　DNA 和 RNA 在复制、转录和翻译中的中心作用，以及 DNA 和 DNA 之间、DNA 和 RNA 之间、RNA 和 RNA 之间的互补作用，启示人们设想能否通过 DNA 片段或 RNA 片段来调节生物信息的复制、转录和翻译过程，从而达到治疗某些疾病的目的。人工合成寡聚核苷酸及其衍生物对肿瘤的治疗已经取得可喜的进展。作为蛋白质翻译抑制剂的寡核苷酸现称为反义 RNA 或反义 DNA，其作用原理可能有以下几种：①反义 RNA 与体内靶 mRNA 结合形成双螺旋，使 mRNA 不能成为翻译蛋白质的模板；②反义 RNA（DNA）与 DNA 结合，影响 DNA 复制；③反义 RNA 能阻止 mRNA 与核蛋白体结合，降低蛋白质合成效率。

第三节　蛋白质和多肽的氨基酸序列测定

蛋白质是生物体的基本组成成分，也是生物体中含量最丰富的生物大分子。测定蛋白质的氨基酸序列是进行蛋白质结构与功能研究必不可少的，其重要性在于：①可以根据某些蛋白质分子的氨基酸顺序在进化中的变异，绘制物种发生树，揭示这些物种之间的进化关系；②蛋白质一级结构的研究是在蛋白质分子水平上阐述某些疾病发生的机理及诊断分子病的基础，如镰刀状贫血是由于遗传基因突变，导致血红蛋白一级结构突变，结果原β链第 6 位的谷氨酸变成了缬氨酸，导致其功能缺陷；③蛋白质一级结构的研究可为多肽类新药的设计与合成提供依据；④蛋白质一级结构的研究是开展蛋白质工程工作的前提，根据蛋白质结构与功能关系的研究，对蛋白质分子中某些氨基酸残基进行修饰或置换，以了解某个或某些氨基酸残基对其功能的作用，并采用基因工程手段对蛋白质分子进行定向改造以获得具有新的特性的蛋白质，即为蛋白质工程；⑤蛋白质一级结构的研究及与已知蛋白质同源性的比较，可为该蛋白的功能研究提供线索。

一、序列测定前的准备

（一）蛋白质纯化

每种蛋白质都有它特定的氨基酸排列顺序，因此作顺序测定的蛋白质必须是纯的，测出的顺序才有意义，通常情况下杂蛋白含量超过 5%时便很难测定顺序。在熟练地掌握测定顺序方法的前提下，关键是如何拿到一个纯的蛋白质样品，然后把这个蛋白质切成若干肽段，再分离提纯。确定蛋白质（或多肽）样品的纯度标准大致如下：

① 分子大小是否均一；

② 电荷是否均一；

③ 是否具有恒定的氨基酸组成；

④ 是否具有单一的 N 末端和 C 末端残基；

⑤ 进一步纯化时生物活性是否再有提高；

⑥ 能否结晶；

⑦ 从 N 末端测几步顺序，一般在 4 步以上，如果都是单一的氨基酸残基，则含杂质可能约为十六万分之一。

上述7条是比较严格的要求，很难全部达到。通常，以电泳时呈现一条带以及末端残基鉴定是单一的即可进行测定顺序。

蛋白质很复杂，往往具微观不均一性，这种不均一性可分为两种情况：一种是在蛋白质肽链合成后在加工中造成的，如糖蛋白，由于含糖量不同，它们的电泳行为常表现很大差异；又如赖氨酸的甲基化、丝氨酸的磷酰化以及磺酰化等，也会造成电泳行为的不均一性；还有一些则是在分离纯化过程中人为造成的，如脱酰胺等。这种不均一性对蛋白质顺序分析不会带来太大的困难。另一种不均一性是由于基因家族产生的，它们之间的不同表现为个别氨基酸的替换、N端或C端肽段的缺失等，这种不均一性会给蛋白质顺序分析带来很大的困难。由于顺序测定工作流程长，工作量大，在顺序测定前如发现样品不纯应该仔细区分，然后再决定相应的对策。

（二）肽链的分离

必须注意，如果发现该蛋白质含有两条或两条以上的多肽链时，首先要设法将它们拆开，然后再分离提纯。多肽链的数目通常可以从测定每分子蛋白质所含的N末端残基数，结合SDS凝胶电泳推导出来，显然两者数目应一致。若肽链之间非共价交联，可用高浓度变性剂（如8mol/L尿素）或解聚剂（如SDS）拆离。若肽链是由二硫键共价交联，就得选用适当的化学试剂断裂二硫键，并将断裂后出现的巯基保护起来，然后再分离提纯。断裂二硫键最常用的方法是过甲酸氧化法和还原-羧甲基化法。

1. 非共价键的分离

肽链间的结合多数以非共价键结合，如氢键、疏水键及静电引力等。非共价键的结合较弱，通常可用酸、碱、脲或胍等变性剂和SDS等去污剂将其破坏。

（1）改变pH法　对于蛋白质的解离临界pH约在3～4和9～10的范围，当加入酸或碱，使蛋白质溶液pH低于3或高于10时，可使肽链分离。这是最简便、实用的方法。

（2）变性剂法　大部分蛋白质在8mol/L脲或6mol/L盐酸胍的溶液中发生变性。此种变性多为蛋白质肽链的解离。但本法不足之处是高浓度变性剂可使很多后来的化学和物理研究复杂化；除去变性剂又可引起重新缔合或无规则聚合。

2. 共价键的分离

肽链间存在共价键结合，须用特殊的化学作用将其破坏。例如，二硫键的切断可用氧化法和还原法。还原法使用较多，系将二硫键还原成巯基，并使二硫链断裂，不足之处是还原剂去除后，还可重新形成二硫键，克服的办法是用羧甲基化封闭巯基，或加大巯基化合物浓度。氧化法的优点是切断二硫键后，肽链不能重新形成二硫键，但有可能破坏其他氨基酸侧键。

3. 肽链的鉴定与分离

蛋白质的多肽链拆开后，须鉴别它们是否相同，如不相同，需设法将它们分离。鉴别的方法有SDS-聚丙烯酰胺凝胶电泳、N末端氨基酸分析等。分离的方法有离子交换层析、凝胶过滤、吸附层析及等电聚焦等，一般解离剂不影响上述方法的分离效果。

（三）肽链的部分裂解

一条多肽链往往过长，给直接测定多肽链的氨基酸顺序带来困难，因此一般在取得纯的单肽链后，须先分解成适当大小的肽片段，才能进行顺序分析。单肽链的分解要求降解点少、选择性强和反应产率高，是顺序分析成功与否的关键性步骤，目前使用的方法可分为两大类，即化学法和酶解法。

1. 化学法

包括酸水解、碱水解及专一性化学试剂如溴化氰裂解等方法。

(1) 溴化氰 (CNBr) 裂解法 在化学法中最为重要和最常使用。它专一降解肽链中的甲硫氨酸，反应专一，产率也较高；又由于蛋白质中的甲硫氨酸的数目较少，所以降解的肽片段也较少，新生成的肽段往往是大片段，适合于自动测序仪测定，更适合于固相法测定。

反应是在酸性介质中进行。在酸性条件下氨基质子化，可以避免一些副反应。如用70%甲酸可以破坏肽链的卷曲，使甲硫氨酸侧链暴露，有利于溴化氰的裂解。

(2) 部分酸水解 在稀酸 (0.03mol/L 盐酸) 的条件下，天冬氨酸后面的肽键容易断裂，如酸浓度较高 (11mol/L 盐酸)，则含羟基的侧链残基氨基侧的肽键容易断裂。酸水解虽然有一定的规律性，但专一性不如其他方法。

(3) 其他 除对甲硫氨酸残基分解法 (CNBr 法) 外，还有半胱氨酸残基和色氨酸残基的专一性分解法。半胱氨酸残基专一分解法是在 pH8 条件下，以 2-硝基-5-硫氰基苯甲酸与肽链的半胱氨酸反应，生成 5-氰基半胱氨酸，残基氨基侧的肽键断裂，获得 N 末端受到封锁的肽段。对于肽链中的色氨酸残基羧基侧的肽键，可用 2-硝基苯亚磺酰肌-3-甲基-3-溴代吲哚氧化，使其断裂。因色氨酸在蛋白质中存在甚少，经此处理可得到较大的片段。

2. 酶解法

它与化学法相比降解的专一性较强，其效率也较高、副反应也少，而且肽片段易纯化，对在酸水解中易受破坏的谷氨酰胺、天冬酰胺及色氨酸等氨基酸在酶降解中不受影响。

酶解之前，对于蛋白质需进行变性处理，使其空间结构破坏，酶易与蛋白质分子作用。变性方法多用加热法或三氯乙酸沉淀法。

几种常用的蛋白水解酶见表 10-1。

表 10-1 蛋白水解酶及其特性

酶	特异性	最适 pH
胰蛋白酶	断裂 Lys 和 Arg 的羧基端	8～9
胰凝乳蛋白酶	断裂 Met、Tyr、Trp、Phe 及 Leu 的羧基端	8～9
嗜热菌蛋白酶	断裂疏水性氨基酸 Leu、Ile、Met、Phe、Trp 及 Val 的氨基端	8
胃蛋白酶	主要作用芳香族氨基酸和 Leu 的羧基端	2
梭菌蛋白酶	断裂 Arg 的羧基端	8
鼠颌下腺蛋白酶	断裂 Arg 的羧基端	8～9
金黄色葡萄球菌蛋白酶	断裂 Glu、Gln、Asp 及 Asn	4
枯草杆菌蛋白酶	专一性很差	8

选择合适的水解酶后，还要具备下列条件：适宜的 pH 和反应温度；恰当的酶与底物的比例和反应时间。在实际工作中，尚无规律可循，需要进行预备试验来确定最佳条件。

胰蛋白酶是最常用的蛋白质水解酶，其专一性强，只断裂赖氨酸 (Lys) 和精氨酸 (Arg) 的羧基端肽键。如被降解的肽链含有较多的 Lys 和 Arg 时，断裂的片段也较多。此时可用柠檬酸酐和马来酸酐等保护 Lys，使胰蛋白酶只裂解 Arg 的羧基端，或者用 1,2-环己二酮保护 Arg，使胰蛋白酶专一水解 Lys 的羧基端。

对专一性较差的酶，如枯草杆菌蛋白酶和木瓜蛋白酶，可采用短时间、低浓度等特定条件进行有限的反应，仅使特别敏感的肽链发生水解而获得较大的片段。

(四) 肽片段的分离

肽片段分离的主要目的是为随后的序列分析提供足够数量的纯净样品。它在一级结构分

析中最为费时和最烦琐，且工作量很大。分离的原理往往是基于肽片段的物理化学特性，如分子大小、电荷、极性、溶解度及特殊的共价或非共价结合性质等。肽的分离方法很多，如凝胶过滤、离子交换层析、薄层层析、高效液相色谱等。在实际工作中，常常将上述方法结合使用，以便达到分离和提纯肽片段的目的。

（1）凝胶过滤　常用于降解后的肽片段的粗分，尤其适用于较大肽段的分离。肽段组分经分离后，往往需要再进行下一步提纯，如过离子交换层析柱。

（2）离子交换层析法　较大的肽段选用离子交换纤维素或离子交换葡聚糖凝胶，对较小的肽段可选用离子交换树脂如 Dowex 50。

（3）薄层层析法　对于溶解度高的小肽（10 个残基以下），效果较好，不适于大片段的分离。

（4）高效液相色谱法（HPLC）　已成为目前最有用的肽片段分离技术之一。已有分离多肽和蛋白质的专用 HPLC，称为快速液相色谱仪（FPLC）。其优点是速度快、灵敏度高、分辨率好、样品用量少（μg 级）。常用反相高压液相色谱法，即用疏水性非极性的柱填料，如表面结合有长链并用烃基涂层的刚性二氧化硅球体，直径小于 10μm，在极性溶液中，其有较强的吸附肽片段的作用。在 20065～50163kPa 的高压下，用甲醇或乙腈等有机溶液洗脱。通过降低洗脱液的极性，使被吸附的肽段按其极性大小顺序从柱上洗脱下来，达到分离肽片段的目的。

（5）肽片段的检测　最方便的方法是直接测定，已知肽键在 200～210nm 处有强吸收。也可用茚三酮反应显色，不足之处是灵敏度偏低。亦可用荧光胺反应测定反应产物的荧光强度（激发 390nm，发射 475nm），其灵敏度高出茚三酮反应 1～2 个数量级。此外，可利用邻苯二醛和巯基乙醇与肽片段反应产生荧光；较荧光胺反应更为灵敏，测定条件为在激发 340nm，发射 455nm 波长下测定反应产物的荧光强度。

二、序列测定的方法和原理

蛋白质（肽段）氨基酸顺序测定既可从 N 末端开始进行，也可从 C 末端开始进行。但是，从 C 末端测序的方法尚不成熟，通常都是从 N 末端进行。分析肽片段氨基酸排列顺序的方法有手工法和自动分析法。自动分析法又有自动液相测序仪、自动固相测序仪及气相测序仪；而手工法常用的有 Edman 降解法、DABITC/PITC 双偶合法、DNS-Edman 降解法及减数 Edman 法等。一般是依据实验室的仪器设备条件、实验操作者的经验与技巧，以及肽的大小和氨基酸组成情况，选择使用上述分析肽片段氨基酸排列顺序的方法。

（一）手工测序

获得纯肽以后，可以开始测定肽段的氨基酸序列。测定方法主要是化学法，酶学方法也有一定意义。

1. Edman 化学降解法

此法迄今仍然是蛋白质或多肽序列测定技术中最有效和最基本的方法。降解所用的试剂是异硫氰酸苯酯（phenylisothiocyanate，PITC）。降解全过程包括三个主要反应。第一步为偶联反应，肽的 α-氨基与 PITC 在碱性条件下发生偶联反应，产生苯异硫甲氨酰肽（简称 PTC-肽）。第二步为断裂反应，在强酸作用下，PTC-肽的近 N 端肽键断裂，产生噻唑啉酮苯胺（thiazolinone，ATZ），与此同时形成一个少一个氨基酸残基的新肽。新肽的 N 端仍是自由的，可以进一步与 PITC 反应，进行第二次降解。第三步为转化反应，断裂产物 ATZ

不稳定，为便于随后的 N 端氨基酸分析测定，要将这个不稳定的化合物转化为 3-苯基-2-乙内酰脲（phenylthiohydantoin-AA，PTH-AA）。整个反应如图 10-7 所示。生成的 PTH-AA 非常稳定，可用薄层层析、气相色谱、回水解法（即用酸水解 PTH-AA，使其转化为游离氨基酸，再用氨基酸分析仪测定）、HPLC 等方法进行测定，即得氨基酸序列的第一个残基。

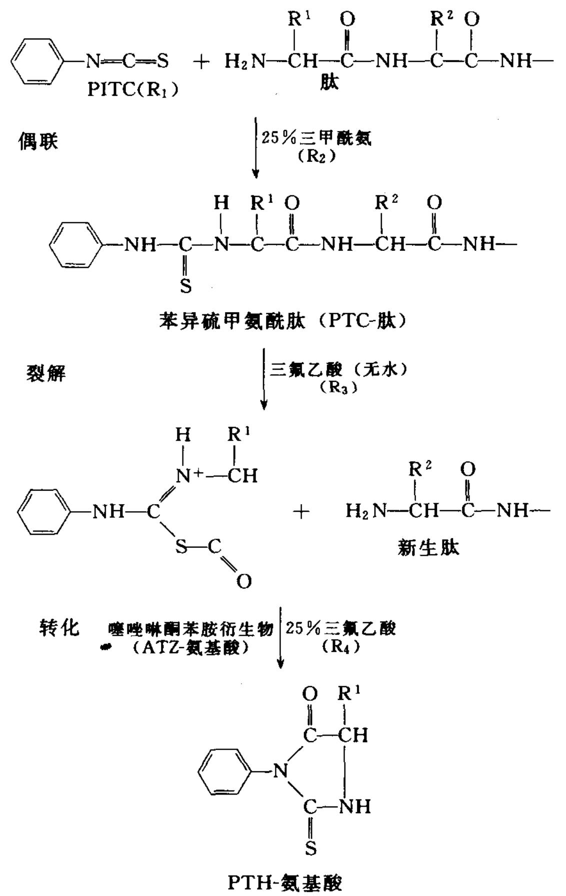


图 10-7 Edman 降解法反应示意图

从上述反应步骤可知，蛋白质或多肽的自由 α-氨基与 PITC 偶联后，与其紧挨的第三个残基的键合力大大减弱，很易断裂，所以在无水酸作用下，仅切下与自由 α-氨基反应的那个残基。当切下 N 端这个残基后，紧挨的那个残基（第二个）便暴露出来，成为自由 α-氨基酸，又可与 PITC 进行偶联反应。然后，再切下第二个残基，以此循环。它给序列测定带来许多方便，成为 Edman 降解反应的最大优点，液相序列仪和固相序列仪及新近问世的气相序列仪就是根据 Edman 降解反应原理设计成的。

利用 Edman 降解一次能连续测出多少个残基，关键在于肽转变成 PTC-肽及随后的 Edman 降解反应产率。假如这两步反应的总收率是 95%，那么到测定第 15 个残基时，从反应开始前的肽段开始计算，反应产率是 46%。反应进行到此，将有前 14 步反应残留的各种东西夹杂，结果难于分析清楚。因此，上述反应的产率是十分重要的因素。使用非常纯的试剂，并在严格控制的条件下进行各步反应，则上述反应的重复收率可高达 98%。在此基础上，可连续测定 60～70 个残基。因此，近年来，肽片段的制备趋向于得到大肽段，因为直接测出大肽段的序列后，其他工作就省事多了。

2. DNS-Edman 序列分析法

这是对 Edman 降解技术的一种改良方法。此法将高灵敏度的 DNS 技术与能连续降解的 Edman 降解反应有机地结合起来，其原理是先取一小部分肽与 DNS-Cl 反应，按末端测定法测出该肽的第一个残基，然后用 Edman 降解剩余肽的第一个残基，经抽提、弃掉，同时得到少一个氨基酸残基的新肽。再取一部分新肽与 DNS-Cl 反应，酸水解后，得到一个具有强荧光的氨基酸衍生物，通过鉴定可知其第二个氨基酸残基，其余部分再进行 Edman 降解，降解后的新肽再取一小部分与 DNS-Cl 反应，经水解后鉴定 DNS-氨基酸，则可知该肽链的第三个氨基酸残基，如此循环下去，就可测出氨基酸的序列。

由于水解后的生成物是带荧光的 DNS-氨基酸，其检出的灵敏度比 Edman 法高几倍甚至十几倍，因此用样量可减少至 1/10，而且又能连续进行，这是此法的优点；缺点是因为要

用酸高温水解，使 Gln 和 Asn 相应地变成 Glu 和 Asp，而且不能检出色氨酸。

3. 减数 Edman 序列分析法

这也是 Edman 降解技术的一种改良方法。此法是利用 Edman 降解法将肽链中的 N 端氨基酸残基逐次从肽链中除去，并随之分析相应肽链的氨基酸组成，从每一循环中氨基酸的丢失或减少，可以推出 N 端氨基酸的排列顺序。

对于小肽来说，此法更快速、准确，只要有一台氨基酸自动分析仪就够了，少于 10～15 个残基的小肽可用本法来测定序列。

4. DABITC/PITC 双耦合法

Chang 等介绍一种改进的 Edman 试剂，4-*N*, *N*-二甲基氨基偶氮苯-4-异硫氰酸酯（DABITC）。它先与肽的 N 端结合，经酸裂解并转化后，N 端氨基酸从肽链上裂解下来，转化成有明显颜色的二甲氨基偶氮苯硫氰酯氨基酸（DABTH-氨基酸）。在聚酰胺薄膜上层析分离 DABTH-氨基酸后，经盐酸蒸气熏即能显示红色斑点，而杂质与试剂的反应物显蓝色或紫色斑点。这种颜色差别大大提高了检测的灵敏度和准确性，只要用 pmol 样品即可鉴定。DABITC 降解肽的原理与 Edman 降解相同。

DABITC 虽有明显优点，但由于它与氨基耦合反应产率较低，一般不到 50%，不能有效定量地去除 N 端残基，从而使肽链的降解步数受到很大限制。这个问题可用双耦法来解决，即将余下的没有与 DABITC 耦合的氨基再用 PITC 作第二次耦合。由于 PITC 与多肽的耦合反应近于定量，从而可定量地去除 N 末端，并暴露第二个氨基，暴露的 N 端氨基又能与 DABITC、PITC 作双耦合降解，这样依次循环下去，逐个测定蛋白质或多肽的 N 端氨基酸序列。目前，此法已成功地用于微量序列分析，无论是用液相还是固相，手工或自动方法测序，都得到较满意结果。

5. 酶学降解测定肽序列

（1）氨肽酶法　外肽酶能从端基逐个往里切，因此，原则上说，只要能跟随酶水解进程，将释放的氨基酸分别定量测出，就能测出肽的序列。

氨肽酶是专一性从 N 端逐个降解蛋白质或多肽的一类外肽酶，其种类很多，最常用的是亮氨酸氨肽酶（LAP），在 N 端的第二个氨基酸是 Pro 时，则此酶不能将 N 端水解下来。除此以外，所有的肽键都能被 LAP 水解，但水解速度变化很大。N 端为极性氨基酸和芳香族氨基酸时，水解速度慢，N 端为疏水侧链较大的氨基酸（如 Leu）时，水解速度最快。因此实际操作时，常会无法分辨哪个氨基酸先下来。

（2）羧肽酶法　用化学法测定羧端比较困难，因此，相比之下羧肽酶法重要一些。常用的羧肽酶有三类：A、B、C。其中，羧肽酶 A 最为常用，但用它也只能确定 4～5 个氨基酸残基序列。

（3）外肽二肽酶测定肽序列　有一类外肽酶能每隔 2 个残基切 1 次。其用法如下：第一次将肽段用此类酶水解，得到一系列二肽，经分离纯化后，测出这些二肽的结构；第二次，先将原来的肽段从 N 端切去一个残基，再用外肽二肽酶水解，得到切口与第一次酶解的切口错开的一系列二肽，再分离纯化，测出这一批二肽的结构。将 2 套二肽相比较，就能得到被测肽的序列。

由于二肽结构测定非常容易，它们的分离也稍为容易一些，所以此法优越性很大，关键在于二肽酶的专一性是否符合要求。

从 N 端连续往里切下二肽的酶称二肽氨肽酶，从 C 端往里连续切下二肽的酶称二肽羧

肽酶。二肽羧肽酶很不稳定。

（二）自动序列分析

自动分析方法的出现，加速了蛋白质顺序测定工作，同时减轻研究人员的繁琐手工劳动。蛋白质测序仪的基本原理就是 Edman 降解法的原理，即逐个降解 N 端残基，然后自动顺序检测出残基性质与含量，最后自动报告肽的氨基酸序列。自动测序仪有液相序列仪、固相序列仪和气相序列仪 3 种。

1. 液相序列仪

液相序列仪主要采用了自旋式反应器，在其中可以有效地进行反应和充分抽提，适宜于较大片段的分析。应用 5nmol 样品，可以连续测定 20～50 个左右的残基，收率为 97%以上。液相序列仪的主要缺点是不适于分析小肽，因为小肽在有机溶剂中溶解度大，易于从反应杯中洗出。若将小肽固定在载体上再进行 Edman 降解，则可克服这个缺点，于是产生了固相序列测定仪。

2. 固相序列仪

1971 年 Laursen 提出用固相法测定蛋白质序列，不久又研制了自动固相肽序列仪，弥补了液相序列仪的某些不足。它对一般的小肽（少于 30 个残基的肽），甚至是疏水性肽都能进行序列分析。固相法是先将肽段共价结合到一种惰性载体上，并装成柱，然后按一般的 Edman 降解程序从 N 端逐级降解。固相载体最初多用聚苯乙烯树脂，近年主要用微孔玻璃球，球表面接有丙氨基侧链，然后在它上面还可连接其他反应基团，以便连接各种不同类型的肽。肽的类型不同（C 末端残基或侧链基团不同），结合到载体上的方法也不同。目前固定肽的方法大体有对苯二异硫氰酸法、高丝氨酸内酯法和水溶性碳二亚胺法。

由于切下一个残基后的肽仍在载体上，当然不会被溶剂洗掉，可连续进行 Edman 降解反应。降解下来的氨基酸与母体肽链，只需通过简单过滤即可完全分离，仪器本身及操作也较简单，容易掌握。固相法对试剂要求通常比液相法低，因此处理简便，价格便宜。固相法测序列的缺点是肽链共价结合到载体上的收率较低，有时低到 20%～30%，所以肽与载体的偶联方法及载体的开发仍是目前要解决的问题。近年来，随着 DABITC/PITC 双耦合法应用的广泛深入，已将它应用于自动固相序列仪中，取得了良好的效果，可以进行微量序列分析。

3. 气相序列仪

自动化蛋白质序列仪刚问世时，需要样品量约 1～2mg。1978 年 Tarr 等将一种四级铵盐聚合物"polybrene"引入了蛋白质、多肽的序列测定，它能和肽链中的极性基团作用（非共价键合）而将蛋白质和多肽有效地固定在玻璃纤维支持膜上，防止了萃取时样品的损失。为了适应分子生物学对蛋白质微量分析要求，20 世纪 80 年代初，Hewick 及 Hunkpillar 等改进了自动化蛋白质序列仪中的仪器结构，它用弹筒型玻璃反应室（内部体积约 0.05ml）代替了液相序列仪中的旋转玻璃杯，用 polybrene 固定样品，Edman 降解反应中有的试剂如三甲胺以气体方式输送，其他试剂以流动或液相脉冲方式输送。此仪器较以前几种序列仪具有以下明显优点：①灵敏度高，耗费样品量少，通常仅需 10～100pmol；②试剂和溶剂消耗少，大约是液相序列仪的 1/10，降低了费用；③缩短了每步降解循环时间，提高了分析效率。

20 世纪 80 年代末又对气相序列仪进行了部分改进，即将原来三氟乙酸以气体方式输送改为液相脉冲方式，称为脉冲液相序列仪，以适应不同样品的分析需要。另外，由于载有活性基团的功能性 PVDF 膜的问世，蛋白质和多肽可以共价固定在此类膜上，直接在上述两

种序列仪上进行固相序列分析。

在这两种序列仪中，Edman降解后的PTH-氨基酸衍生物可及时直接从转化腔中自动输入高效液相色谱仪中进行PTH-氨基酸衍生物的定性及定量分析，不仅快速可靠，而且防止了样品的损失。两台仪器统一用电脑控制，包括化学反应和分析鉴定以及数据的自动记录和处理。

三、影响测序的因素

1. 影响N端测序的因素

在测序中往往可以发现，即使N端很均一的蛋白质（第一个循环出现单个氨基酸残基），经过十几个循环后背景明显不及第一个残基清晰，这是由于Edman反应是一个化学过程，在其偶联、裂解、转化等步骤都有可能发生一些副反应，从而影响最终裂解效率。目前最佳产率为95%～98%，因此经过50次循环后，PTH-氨基酸分析谱中将出现较多杂峰，影响正确辨认，也就是说对于一般蛋白质最多能分析至N端50个氨基酸左右，欲知全顺序，则需将蛋白质降解成一系列肽段分析后再拼接。

对于有些蛋白质由于含有较多的对Edman反应敏感的残基或肽键，裂解效率将更低。

有时循环至Ser和Thr时产率突然降低，这是因为在进行这两个氨基酸残基分析前一个循环中三氟乙酸除起裂解作用外，可能与Ser和Thr上的羟基反应，继而部分封闭Ser和Thr的N端α-氨基。

2. 影响C端测序的因素

C端测序副反应较多，最高初始产率仅为20%，而对于Glu、Asp、Ser、Thr等氨基酸，其产率将更低。如C端富含上述几种氨基酸，测序将十分困难。另外，一旦遇到Pro，由于吡咯环的存在而不能与乙酸酐反应成环，整个测序过程将终止。

到目前为止，C端序列可测1～10个残基，平均3～5个，一次测序需要的量比N端测序要大得多，至少需1nmol。此外，因为不同的氨基酸的反应产率不同，所以图谱的分析也比N端测序复杂，需要较多的经验。目前C端测序结果最好的一个样品是马肌红蛋白原，可测C端10个以上残基，通常将其作为标准样品来判断仪器的稳定性。

目前，C端序列测定技术已初步成熟，并开始发挥作用，但其反应的效率与N端Edman降解测序法有一定的差距。因此，目前的主要问题是如何提高反应的产率，特别是如何改善T、S、D、E几种氨基酸的测定，以及如何解决Pro终止测序的问题。如果在上述几方面取得进展，将使C端可测残基数大大增加，并减少所需的样品量。另外，近年来发现很多活性多肽的C端是酰胺化的，如果能检测酰胺化的C端残基，将对活性多肽的研究有很大的帮助。

C端测序技术仍处于研究发展阶段，虽然该技术还不尽完善，但对于确认表达蛋白的C端是否正确，以及判断在表达、纯化过程中是否发生了不必要的加工大有帮助。

第四节　蛋白质或多肽的化学合成

从化学结构上看，蛋白质和多肽都是由氨基酸组成的。各个氨基酸之间通过一个氨基酸的羧基同另一个氨基酸的氨基以酰胺键（也叫肽键）结合连接成肽链，由2个氨基酸连成的

叫二肽，三个氨基酸连成的叫三肽，并依次类推。

化学合成蛋白质和多肽的任务就是如何把各种氨基酸按照天然产物的氨基酸排列顺序和连接方式连接起来。多肽的合成主要有液相合成法和固相合成法。在液相肽合成中，每次接肽以后都需要对产物进行分离纯化或结晶，以除去未反应的原料和副产物，这个过程是一个高劳动强度的过程，往往需要花费几个月的时间而总收率却极低。为了减少这些繁琐的工作，并期望能够使接肽反应自动化，Merrifield 于 1963 年发展成功了固相多肽合成方法。固相多肽合成的前提是由氨基酸组成的肽链的一端挂在一个不溶性的高分子载体上，将肽链连接在高分子载体上的最大优点是合成中所有的纯化步骤都被省略了，而以简单的冲洗和过滤肽树脂代之。因此，液相合成中由于转移过程带来的损失和中间纯化步骤的损失消除了。简言之，固相法的优点在于简化了操作，减少了损失，所有反应及操作均在一个反应管中进行，克服了用液相法合成长肽分子时在溶液中难以溶解的问题。

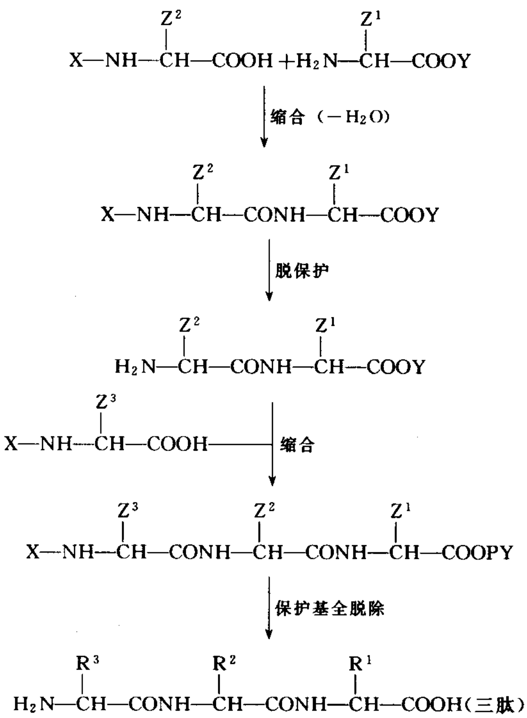


图 10-8　多肽合成原理

其中 X、Y 分别为 α-氨基和 α-羧基保护基，

Z^1、Z^2、Z^3 分别为侧链基团

R^1、R^2、R^3 的保护基

固相合成法由于其自身的许多优点而得以广泛应用。无论是液相合成还是固相合成，从原理上来讲并不复杂，大体上可分为三个步骤：①官能团的保护；②羧基的活化和肽键的生成；③合成肽的脱保护基和纯化。一般可表示成如图 10-8。

一、氨基酸官能团的保护

为使肽键生成反应定向进行，需要对氨基酸的各种官能团加以适当的保护。自从 Bergmann 首先用 Z 基团成功地保护氨基以来，已出现各种各样的保护基，但在肽合成中常用保护基仍然是相当限定的。

1. α-氨基的保护

常用的氨基保护基可分为烷氧羰基、酰基和烷基。其中，烷基羰基保护基可防止消旋化，因此常被采用为 α-氨基保护基，如叔丁氧羰基（Boc）和芴甲氧羰基（Fmoc）等。

2. α-羧基的保护

α-羧基一般以盐或酯的形式加以保护。盐是对羧基的临时性保护，常用的有钾盐、钠盐、三乙胺盐、三丁胺盐等。常用的酯类有甲酯和乙酯、苄酯、叔丁酯；叔丁酯是近年来最常用的羧基保护基，该基团可用酸在温和条件下脱除。

3. 侧链官能团的保护

为了避免副反应的发生，某些氨基酸的侧链官能团需采用适当的保护基加以保护。采用什么种类的保护基要根据合成以及脱保护条件而定。在 20 种氨基酸中有 13 种氨基酸侧链具

有活泼官能团，在多肽合成时往往需要保护。当然，有些基团是在任何情况下都必须保护的，有些基团则依情况而定。

（1）赖氨酸的ε-氨基保护　赖氨酸的ε-氨基是必须加以保护的。原则上，α-氨基的所有保护基都可作为ε-氨基的保护基。但ε-氨基的保护与α-氨基保护不同，该保护基往往在肽链合成之后才除去，因此要选择具有与α-氨基保护基不同脱除方式的保护基。此外，侧链ε-氨基的保护没有消旋问题，可以采用酰基型保护基。侧链ε-氨基保护基常采用Z、Boc、Fmoc等，一定要注意与α-氨基保护基在脱除方式上搭配使用。

（2）侧链羧基的保护　如果C端残基的α-羧基是通过酰肼和叠氮物活化，那么这个羧基组分的侧链羧基就不用保护。但在一般情况下，天冬氨酸和谷氨酸的侧链羧基总是需要保护，常采用苄酯和叔丁酯保护。在引入叔丁基保护之前，要对天冬氨酸的α-氨基和羧基保护（Z，OBzl)，然后引入侧链保护基团。叔丁基可用三氟醋酸（TFA)、三甲基溴化硅烷(TMSBr)、三氟甲磺酸硅烷酯（TMSOTf）等脱除。用催化氢化法脱除苄基后，可得到叔丁基保护的天冬氨酸。值得注意的是，以该衍生物合成肽时，其产物在酸、碱条件下很容易环化，生成琥珀酰亚胺副产物，尤其当天冬氨酸的羧基与丝氨酸、苏氨酸和甘氨酸邻近时，即使在中性条件下，也很容易产生这种副反应，近年来，可以减轻这种副反应的几种新的保护基β-环烷醇酯、金刚烷醇酯和薄荷醇酯逐渐推广起来，这些保护基可用TMSOTf除去。

（3）胍基的保护　精氨酸的胍基有很强的碱性和亲核性，很容易被酰化，所以必须加以保护。理想的情况是精氨酸胍基的3个氮原子都应得到保护，但实际上常常只保护1个或2个胍基氮。

常用的胍基保护基分为硝基、烷氧羰基、磺酰基和三苯甲基4类。近年来，4-甲氧-2,3,6-三甲基苯磺酰基（Mtr）受人欢迎。Mtr基团用TFA/thioanisole 30min即可脱去，三苯甲基（Trt）也被用来保护胍基的ω-氮上，用酸可以脱除。

（4）半胱氨酸的巯基保护　半胱氨酸的巯基有很强的亲核性，容易与氨基竞争被酰化成硫醚，也容易被氧化成二硫键，因此该基团是必须保护的基团。常用的保护基分为三类：一是用TFA、$(CF_3COO)_3Tl$/TFA（三氟乙酸铊/TFA）可脱除的基团，包括对甲基苄基(PMeBzl)、对甲氧苄基（MBzl)、Trt等；二是可用$(CF_3COO)_3Tl$/TFA脱除，但对TFA稳定，如叔丁基（^tBu)、乙酰胺甲基（Acm)、苄氧甲基（Bom)；Cys（^tBu）也可用TFMSA和Hg^{2+}脱除；Cys(Acm）可用Hg^{2+}、I和三氟乙酸银（AgOTf）脱除，Cys(Bom）可用AgOTf和TMSOTf/thioanisole脱除；三是对弱酸稳定的基团，有苄基（Bzl）和叔丁硫基（S^tBu）等。Cys（S^tBu）的脱保护是用巯基试剂和含磷试剂等进行还原，而Cys(Bzl）用Na/液氨脱保护基。

（5）丝氨酸和苏氨酸的保护　丝氨酸和苏氨酸不加以保护直接应用于肽合成工作中，虽然也有些成功的报道，但一般情况下需要对其侧链羟基进行保护。常用的保护基为叔丁基和苄基，与天冬氨酸和谷氨酸的情况大体类同，叔丁基的引入比较麻烦。首先，制成苄氧羰基的丝氨酸或苏氨酸的对硝基苄酯，然后在酸催化下同异丁烯反应。苄基的引入用苄醇或溴化苄方法进行。

（6）酪氨酸的保护　与丝氨酸和苏氨酸相同，酪氨酸的酚羟基也可以用叔丁基保护。此外，也可采用对TFA稳定的一些基团，如苄基、2-Br-Z、2,6-di-Cl-Bzl等，但在用强酸脱保护基时，这些基团可能会使酚烷基化。

（7）谷氨酰胺和天冬酰胺的保护　谷氨酰胺和天冬酰胺在肽合成中一般不加以保护。然

而在合成大肽时，要考虑几个不利因素。N-保护的天冬酰胺和谷氨酰胺的 α-羧基被活化时可能产生分子内的脱氢反应，生成氰基化合物。N-保护的天冬酰胺比 N-保护谷氨酰胺更易发生这个反应。在碱性条件下（如 Fmoc 脱保护基时所用的条件），谷氨酰胺的侧链可以环化生成亚酰胺或 α，γ-谷氨酸肽。此外，侧链不保护的 Fmoc-Asn、Fmoc-Gln 在二氯甲烷（DCM）中溶解度较差。因此，为避免这些问题，谷氨酰胺和天冬酰胺的侧链可用 2，4，6-三甲氧苄基（Tmob），4，4′-二甲氧二苯甲基（Mbh）或 Trt 等基团保护。这 3 种基团均对碱稳定，可用 TFA 脱除。Tmob 比 Mbh 更容易用 TFA 脱除。

此外，色氨酸、组氨酸和甲硫氨酸的侧链都需要保护。活化的组氨酸衍生物是最容易发生消旋的氨基酸，其消旋机制主要是分子内的咪唑基的 π-氮原子夺取 α-碳上的氢原子。此反应是碱催化进行的，因此降低咪唑环的碱性可以抑制消旋化。直接保护 π-氮原子的 Bom、叔丁氧甲基（Bum）抑制消旋效果最好。

二、羧基的活化和肽键的生成

在多肽合成中，具有游离羧基的组分被称为羧基组分，具有游离氨基的组分被称为氨基组分，肽键的生成主要涉及羧基的活化，羧基活化的基本原理是将 N-保护氨基酸或肽的 α-羧基转变成活化型的 RCOX，使羰基碳原子带有较强的正电性而有利于 R—$\ddot{N}H_2$ 对它进行亲核反应，取代基团的电负性越强，对羧基的活化能力也越强。比如，X 为卤素时，其衍生物具有很强的反应活性。但是，有时太强的活性会带来不必要的副反应，使应用受到一定限制。因此，经典的酰氯法在形成长链肽时已不采用了。目前常用的肽链缩合法是比较温和的活化酯法、混合酸酐法、叠氮法和缩合剂法。

$$\mathrm{RCOOH} \longrightarrow \mathrm{RCO{-}X} \xrightarrow{\mathrm{R'\text{-}\ddot{N}H_2}} \mathrm{RCONHR'} + \mathrm{HX}$$

1. 混合酸酐法

混合酸酐法是 20 世纪 50 年代发展起来的，方法简单，反应很快，纯度很好，尤其适合于小肽的合成。N-保护氨基酸在叔胺存在下与氯甲酸酯（乙酯或异丁酯）反应生成混合酸酐，再与氨基组分反应生成肽键。

2. 活化酯法

活化酯的概念最早是 1951 年由 Wieland 提出的。他将羧基组分采用硫代苯酯活化形式与氨基组分反应，形成了肽键。但直到 1955 年 Bodanszky 发展了对硝基苯酯后，该法才显示出重要性。Bodanszky 用 N-保护氨基酸的对硝基苯酯同氨基组分反应，逐个增长肽链，成功地合成了催产素。活化酯法具有如下特征：①与混合酸酐法相比，活化酯法的中间体可以分离、结晶、保存；②氨基组分的末端羧基可以不保护；③活化酯法对谷酰胺、天冬酰胺特别有效，对这些氨基酸而言，当羧基活化时，会产生侧链酰胺的脱水反应，在活化酯阶段，可以通过重结晶步骤除去副产物。

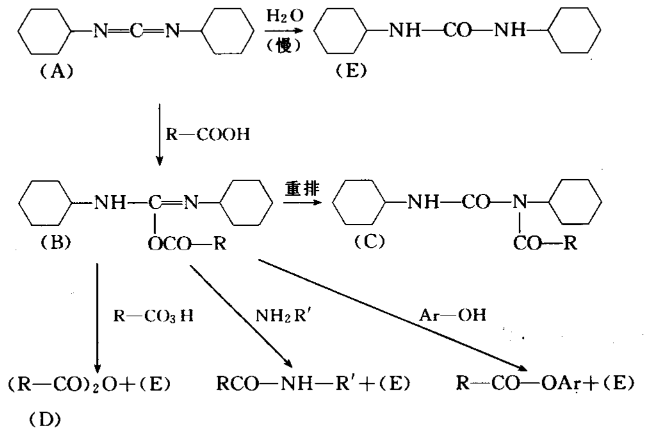
需要注意的是这些酚类活化酯在缩合反应之后生成的游离酚有时很难除净，而 N-羟胺类活化酯与之不同，反应之后很容易除去，因此在肽合成中尤其是液相合成中应用十分广泛。

通常采用 N，N′-二环乙基碳二亚胺（DCC）制备活化酯，即 N-保护氨基酸与醇组分在

DCC 作用下生成活化酯形式。

3. 缩合剂法

缩合剂的种类很多，最重要的是 1955 年 Sheehan 等人提出的 DCC。DCC 接肽反应机制如下。



N-保护的氨基酸与 DCC 反应首先生成 *O*-酰基脲（B），该中间体实际上是一种活化酯，可以直接和氨基成分反应，或者和另一分子酸生成对称酸酐（D），然后和氨基成分反应成肽及 *N*,*N*′-二环己基脲（E）（DCU）。DCU 在大部分有机溶剂中溶解度很小，但过滤有时很难除尽，若采用水溶性碳二亚胺则可避免这个麻烦。近年来，尤其在固相多肽合成中，越来越多地采用二异丙基碳二亚胺（DIC），叔丁甲基碳二亚胺和叔丁乙基碳二亚胺等作为缩合剂，生成的脲可溶于 DCM 中，经洗涤后可除去。

在 DCC 反应中，加入某些添加剂不仅可减少副反应，而且可提高产率。常用的添加剂有 *N*-羧基琥珀酰亚胺（HOSu）、1-羧基苯骈三氮唑（HOBt）和 3-羧基-4-氢-3,4-二氢-1,2,3-苯骈三氮嗪（HOObt）。

HOBt 与 *O*-酰基脲很快反应生成 RCOObt 活化酯，因而抑制了 *N*-酰基气脲形成和消旋化等副反应。HOObt 抑制消旋化效果最好，越来越多地受到重视。

4. 叠氮法

Curtius 于 1902 年建立的叠氮法是最古老的经典肽键生成法，一直沿用至今。其主要原因之一是该法最少引起消旋，因此尤其适用于液相法中肽片段的缩合。

5. 原位法

直接加碳二亚胺和 *N*-*α*-保护氨基酸到树脂中进行反应的方法称为原位法。在 Boc 固相法和 Fmoc 固相法中都很常用，尤其用 DIC 代替 DCC 效果更好。

三、合成肽的脱保护基及纯化

合成肽的最终脱保护基不仅决定合成的成败，而且也决定了侧链保护基的选择等合成计划的确立，因此是肽合成中极重要的环节。目前最常用的脱保护基系统是 TFA、HF、有机磺酸（TFMSA，MSA），以及含硅试剂 TMSOTt、TMSBr 等。

1. TFA 法

TFA 可以脱除一些不耐酸的保护基，如 Bu、Boc、金刚烷氧羰基（Adoc）、Trt 等。本法比较温和，副反应少，越来越多地得到应用，尤其广泛应用于固相合成法中。Wünsch 等

曾用 TFA 最终脱保护法合成了胰高血糖素（glucagon）。在此合成中，侧链保护基采用了ᵗBu、Boc 和 Adoc。在引入含硫氨基酸（Cys，Met）之前，采用 Z 基作为 α-氨基保护基，其后，采用 2-硝基苯磺酰（Nps）、金刚烷氧羰基（Adoc）保护 α-氨基。最终，用 TFA 一次脱去了全部保护基。但 TFA 脱保护法仍有些不足之处，如需要几种氨基酸保护基、Arg 的侧链保护基很难脱除、TFA 用量较大等。另外，TFA 是一个危险的腐蚀性液体，在进行操作时要在通风橱中进行，戴上防护眼镜及手套。

2. HF 法

1967 年，木神原等人报道用无水氟化氢在 0～20℃处理 30～60min 可脱除许多其他方法难于脱除的保护基。Kimura 等用 HF 法脱保护成功地合成了 84 个氨基酸的人甲状腺激素。可见，HF 比 TFA 的脱保护作用更为有效，但不利之处也十分明显，如脱保护基时会产生一些副反应（酪氨酸芳香核的烷基化，谷氨酸的茴香酰化等），可通过控制反应温度、添加适当的捕获剂等方法加以控制。令人讨厌的是 HF 气体剧毒，有很强的腐蚀性，能迅速地溶解玻璃并放热，只能在氟化树脂等特殊反应容器中应用，而且操作十分复杂。操作中一定要避免接触任何 HF 液体或气体，实验应在通风很好的通风橱中进行。吸入 HF 会导致死亡，操作中一直要戴上防护眼镜和手套。

3. 有机磺酸法

考虑到 TFA 法和 HF 法的长短之处，1974 年矢岛和藤井建立了三氟甲磺酸（TFMSA）和甲磺酸（MSA）法。MSA 和 TFMSA 的酸性分别比 TFA 强 16 倍和 426 倍，但二者的氧化性和脱水性都比较弱；它们没有 HF 的强腐蚀性，不需要特殊装置，操作简便。采用 MSA 或 1mol TFMSA-thioanisole/TFA 体系，像 HF 法一样，可在短时间内定量除去 Z、Bzl 等苄醇系保护基，虽然在低温、短时间内除去精氨酸的 Tos 基有一定困难，却可以非常有效地除去精氨酸的 Mbs 和 Mts 等基团。矢岛和藤井用 MSA 和 TFMSA 脱保护法成功地合成了牛胰 RNase A。TFMSA 是最强的酸之一，使用时一定要小心，一定要带上防护眼镜和手套，并在有良好通风的通风橱中进行。

4. 含硅试剂法

继 MSA 和 TFMSA 法之后，1987 年藤井等建立了含硅试剂——TMSOTr 和 TMSBr 两种脱保护方法。在此反应中，TMSOTr 或 TMSBr 作为硬酸（hard acid），thioamsole 作为软碱（soft base），通过推-拉的反应机制使保护基脱除。已证明 TMSOTr-thioanisole/TFA 法比 TFMSA-thioanisole/TFA 法裂解速度快、脱保护效率高，可脱除目前肽合成中常用的各种保护基。

TMSBr 虽然不如 TMSOTr 有效，但对 Met(O) 有很强的还原能力，而且可以减压除去，因此和 TMSOTr 搭配起来，二步脱保护，可以达到较完美的结果。

5. 二硫键的生成

在蛋白质合成中，最困难的问题是二硫键生成反应。蛋白质的二硫键在生物合成过程中可能是在前体原阶段形成的。在化学合成时，多年来一直采用空气氧化法/碘氧化法。空气氧化法会产生大量聚合体，碘氧化法在甲硫氨酸、组氨酸、酪氨酸、色氨酸存在时副反应很严重，因此收率都十分低。1987 年，矢岛、藤井等建立了两个新的方法——$(CF_3COO)_3Tl$ 氧化法和亚砜介导法。

Cys 亚砜和 CysR'd 在酸处理下，生成的 SH 基作用于亚砜的 S 原子，在亚砜的位置生成二硫键。采用 Cys(Acm)(O) 和 Cys(MBzl) 时，用 TFA 处理可生成二硫键，采用 Cys

(MBzl)(O) 和 Cys(MBzl) 时，用 TFMSA 等更强的酸处理可得到二硫键，换言之，用不同的 Cys 亚砜和酸处理方法，可在需要的位置定点生成不同的二硫键。此方法的建立给多个二硫键蛋白质的合成带来了希望。

6. 合成肽的纯化

脱保护基后的肽是一个纯度不好的粗产品，一是在合成肽过程中各种副反应、消旋化等造成的副反应肽，二是在脱保护基过程中由于保护基残留，肽键的断裂、烷基化，Trp 和 Met 等氨基酸的氧化等。因此，根据对目的肽的纯度要求，可采用适当的方法进行纯化。合成的肽纯化方法与天然肽纯化时采用的方法相同，即主要采用凝胶过滤、离子交换层析、反相 HPLC 等方法。

凝胶过滤对于小肽的纯化非常有用，对于大一些的肽有很多方法可以选择，经过第一步的凝胶过滤可以得到分子量相近的部分。但是必须注意的是肽与肽和肽与凝胶的相互作用会导致多肽不按照分子量的顺序流出，这种情况可以通过使用强的醋酸溶液或流动相中加入吡啶来克服。对于合成的大肽要选择适当的方法纯化，如透析、HPLC 或亲和层析等。HPLC 在确定多肽的单一性方面非常有用，需要强调的是用等比例的流动相（恒流）得到的单峰不能证明产物的单一性。

7. 合成肽纯度鉴定

需要牢记的是合成肽的纯度鉴定仅靠一种方法是不可靠的，往往需要综合几种方法所得结果才可判定。常用的方法包括：薄层层析（TLC）；氨基酸组成分析（同时用酸水解和酶水解进行，以判定合成肽消旋程度）；等电聚集电泳；分析型 HPLC；质谱分析；元素分析等。

（戴五星）

第十一章　基因组学技术

基因组（genome）一词是 1920 年 Winkles 从 genes 和 chromosomes 两个词组合而成，用于描述生物的染色体组成和全部基因。

1986 年美国科学家 Thomas Roderick 提出了基因组学（genomics）这一概念，指对所有基因进行基因组作图（包括遗传图谱、物理图谱、转录本图谱）、核苷酸序列分析、基因定位和基因功能分析的一门科学。因此，基因组研究应该包括两方面的内容：①以全基因组测序为目标的结构基因组学（structural genomics）；②以基因功能鉴定为目标的功能基因组学（functional genomics），又被称为后基因组（postgenome）研究。结构基因组学代表基因组分析的早期阶段，以建立生物体高分辨率遗传、物理和转录图谱为主。功能基因组学代表基因分析的新阶段，是利用结构基因组学提供的信息相应地研究基因功能，它以高通量、大规模实验方法及统计计算机分析为特征。2004 年 5 月 22 日基因组学与世界卫生第 57 届会议对基因组学作了如下定义：基因组学是关于基因和它们的功能以及相关技术的研究。基因组学技术包含的范围很广，如全基因组扫描和测序技术、DNA 序列变异检测、核型分析、染色体荧光原位杂交、比较基因组杂交、大规模 SNP（single nuleotide polymorphysm）检测与分析、表达序列标签（expressed sequence tags，EST）作图、差异显示技术、基因芯片、转录组和蛋白质组、大规模转基因和基因剔除以及生物信息学与功能基因组学技术，而且随着研究进程的发展，新技术层出不穷。本章将重点介绍一些有关基因组学的常规技术和新技术。

第一节　结构基因组学及常用研究技术

结构基因组学（structural genomics）是基因组学的一个重要组成部分和研究领域，它是一门通过基因作图、核苷酸序列分析，进而确定基因组成、基因定位的科学。它以全基因组测序为目标，旨在弄清基因组中全部基因的位置和结构，为基因功能的研究奠定基础。结构基因组学的目的是建立高分辨率的遗传图谱、物理图谱、转录图谱和序列图谱；其常用研究方法有脉冲场凝胶电泳（PFGE）、毛细管电泳、基因芯片技术、全基因组随机测序等。

一、基因组作图

染色体不能直接用来测序，必须将基因组这一巨大的研究对象进行分解，使之成为较易操作的小的结构区域，分解后还要将其组装定位在染色体上，才能完整地了解人类基因组全貌，这种分解、组装、定位的过程就是基因作图。根据使用的标志和手段不同，作图有三种类型，即构建生物体基因组高分辨率的遗传图、物理图谱、转录本图和序列图谱。

1. 遗传图谱（连锁图谱）

通过遗传重组所得到的基因线性排列图称为遗传连锁图。它是通过计算连锁的遗传标志之间的重组频率，确定它们的相对距离，一般用厘摩（cM，即每次减数分裂的重组频率为1%）来表示。绘制遗传连锁图的方法有很多，经典的遗传标记（第1代标记）最初主要是利用蛋白质或免疫学等的标记，20世纪70年代中后期建立起来的限制性片段长度多态性（RFLP）、随机引物扩增基因组DNA（RAPD）、扩增片段长度多态性（AFLP）方法成为第1代的DNA标记，这类标记在整个基因组中确定的位点数目可达10^5个以上。第2代标记为可变数量串联重复序列（variable number tandem repeat，VNTR），包括微卫星、小卫星（microsatellite/minisatellite）或短串联重复（short tandem repeat，STR或short sequence length polymorphysm，SSLP）标记、DNA遗传多态性分析等。第3代标记是一类称作SNP的遗传标记系统，即单核苷酸多态性标记。在Internet上的GDB（geneome database）网页上可以方便地查找到迄今已发表的各种遗传标记（http：//gdbwww. gdb. org）。

2. 物理图谱

物理图谱是指DNA序列上两点间的实际距离，它是利用限制性内切酶将染色体切成片段，再根据重叠序列把片段连接成染色体，确定遗传标志之间物理距离［碱基对（bp）、千碱基（kb）或兆碱基（Mb）］的图谱。以人类基因组物理图谱为例，它包括两层含义，一是获得分布于整个基因组30000个序列标签位点（sequence-tagged site，STS，其定义是染色体定位明确且可用PCR扩增的单拷贝序列）。将获得的目的基因的cDNA克隆进行测序；确定两端的cDNA序列，约200bp，设计合成引物，并分别利用cDNA和基因组DNA作模板扩增；比较并纯化特异带；利用STS制备放射性探针与基因组进行原位杂交，使每个100kb就有一个标志；二是在此基础上构建覆盖每条染色体的大片段：首先是数百kb的YAC（酵母人工染色体），对YAC进行作图；得到重叠的YAC连续克隆系，被称为低精度物理作图；然后在几十个kb的DNA片段水平上进行，将YAC随机切割后装入黏粒，黏粒的作图称为高精度物理作图。

表11-1列举了基因组物理图谱数据库的部分相关网站，其中的数据主要来自人类基因组，但同时也包含了其他的一些生物体。

表11-1 基因组物理图谱数据库的部分相关网站

数据库	网址
图谱相关资料	
STS数据库(dbSTS)	http://www.ncbi.nlm.nih.gov/dbSTS/index.html
EST数据库(dbEST)	http://www.ncbi.nlm.nih.gov/dbEST/index.html
CpG数据库(CpG island database)	http://biomaster.bio.no/CpGdb.html
细胞遗传图谱	
GDG	http://gdbwww.gdb.org
MGD	http://www.informatics.jax.org/mgol.html
辐射杂种图谱	
RHdb	http://www.ebi.ac.uk/RHdb
克隆重叠图谱	
YAC克隆	
CEPH-GENETHON整合图谱	http://www.cephb.fr/ceph-genethon-map.html
STS/YAC MAP	http://www-genome.wi.mit.edu/
BAC和PAC克隆	http://www.tree.caltech.edu/
黏粒克隆	http://gea.lif.icnet.uk/
整合图谱	http://cedar.genetics.soton.ac.uk/public-html

3. 转录图谱（表达图谱）

利用EST为位标，根据转录顺序的位置和距离所构建的分子遗传图谱被称为转录图谱。它是染色体DNA某一区域内所有可转录序列的分布图，是基因图的雏形。通过从cDNA文库中随机挑取的克隆进行测序所获得的部分cDNA的5′端或3′端序列称为表达序列标签（EST），一般长300～500bp左右。一般说，mRNA的3′端非翻译区（3′-UTR）是代表每个基因的比较特异的序列，将对应于3′-UTR的EST序列进行RH定位，即可构成由基因组成的STS图。这些EST不仅位基因组遗传图谱的了构建提供了大量的分子标记，而且来自不同组织和器官的EST也为基因的功能研究提供了有价值的信息。此外，EST计划还为基因的鉴定提供了候选基因（candidates）。其不足之处在于通过随机测序有时难以获得那些低丰度表达的基因和那些在特殊环境条件下（如生物胁迫和非生物胁迫）诱导表达的基因。因此，为了弥补EST计划的不足，必须开展基因组测序。通过分析基因组序列能够获得基因组结构的完整信息，如基因在染色体上的排列顺序、基因间的间隔区结构、启动子的结构以及内含子的分布等。

4. 序列图谱（分子水平的物理图谱）

序列图谱是以某一染色体上所含的全部碱基顺序绘制的图谱，既包括可转录序列，也包括非转录序列，是转录序列、调节序列和功能未知序列的总和。其中最粗糙的图谱是遗传图谱，它根据相邻标记（如基因和多态片段）间的重组率来测量相互间的距离；具有1～2Mb长度的限制性酶切片段可被分离和构建物理图谱；YAC等长度在40～400kb的插入片段排列构建高分辨率物理图谱；碱基序列为最高分辨率物理图谱。

二、新基因的分离——cDNA末端快速扩增（RACE）技术

随着分子生物学的不断发展和人类基因组计划的深入研究，对基因的结构、功能及表达的研究越来越成为后基因组计划中的研究热点，而这些研究都是以分离和克隆目的基因为基础。

近年发展了cDNA末端快速扩增（rapid amplification of cDNA ends，RACE）技术。RACE是一种应用特异引物和通用引物从低丰度转录本中快速扩增已知cDNA片段旁侧5′和3′端未知序列的简单而有效的方法，具有快捷、方便、高效等优点，可同时获得多个转录本。因此，RACE技术已逐渐取代了经典的cDNA文库筛选技术，成为克隆全长cDNA序列的常用手段。

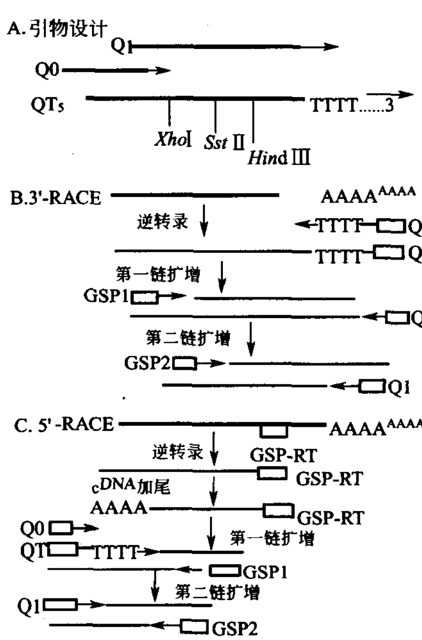


图11-1　RACE原理和方法

1. RACE原理

RACE技术是由PCR技术发展而来，获得cDNA5′和3′末端的方法，也被称为锚定PCR（anchored-PCR）或单边PCR（one-sided PCR）。它是利用oligo（dT）对mRNA进行逆转录的同时在两头加上通用接头（引物），从而可利用基因特异的引物（gene specific primer，GSP）通过PCR反应快速获得目的序列的5′端和3′端。

RACE的引物设计包括锚定引物和基因特异引物（图11-1）。在3′-RACE中，Q0、Q1、QT为3个锚定引物；Q1引物含有编码*Hind*Ⅲ、*Sst*Ⅱ及

Xho Ⅰ的识别位点；Q0 与 Q1 通过一个核苷酸重叠，两者组合后再加上 17 个 oligo-(dT) 序列就构成了 QT 引物(5′-Q0-Q1-TTTT-3′)；GSP1、GSP2 为两个基因特异引物，是根据已知的 cDNA 部分序列设计而成。3′-RACE 的过程如下，以 QT 逆转录总 RNA 得到第一条 cDNA；再用 Q0 与 GSP1 以第一 cDNA 为模板进行第一轮 PCR 扩增，得到双链 cDNA，最后用“嵌套”引物（Q1 与 GSP2）进行第二轮 PCR 扩增，防止产生非特异性扩增产物。5′-RACE 与 3′-RACE 略有不同，首先引物多设计了一个用于逆转录的基因特异引物 GSP-RT；其次在酶促反应中，增加了逆转录和加尾步骤，即先用 GSP-RT 逆转录 mRNA 得到第一条 cDNA 后，用脱氧核糖核酸末端转移酶（TdT）和 dATP 在 cDNA5′端加 poly（A）尾，再用 QT 引物合成 cDNA 第二条链；余同 3′-RACE 过程，用 Q0 与 GSP1（GSP-RT 上游引物）扩增 cDNA，最后用“嵌套”引物（Q1 与 GSP2）进行第二轮 PCR 扩增，以提高特异性。3′-RACE 与 5′-RACE 扩增得到的双链 cDNA 用限制性内切酶酶切，再以 Southern 印迹分析检测酶切后的 DNA 片段中是否存在与已知 cDNA 部分序列制作的探针同源的序列。对同源序列进行克隆即得到 5′和 3′特异性的 cDNA 产物。最后从两个有相互重叠序列的 3′-和 5′-RACE 产物可获得全长 cDNA；或者通过分析 RACE 产物的 3′端和 5′端序列，再合成相应的引物，扩增出全长 cDNA。

2. 技术要点

RACE 技术是延伸已知部分外显子序列和克隆全长 cDNA 基因的主要方法之一，广泛用于许多已知功能基因片段的进一步延伸和全长 cDNA 的克隆，但由于其过程复杂、涉及多步连续的酶促过程，如逆转录、TdT 加尾、第二链合成、RACE-PCR 扩增及 RACE 产物克隆等，在实际应用中有许多问题和相当难度。因此，许多研究小组对 RACE 的各个步骤都进行了改良和优化。

（1）逆转录　影响 RACE 结果的最关键步骤是 mRNA 逆转录合成第一链 cDNA。这在克隆全 mRNA 的 5′端时尤其突出，因为不完全的 cDNA 仍然可以用 TdT 加尾和随后被 PCR 扩增。这些不完全分子与全长分子具有相同的末端，它们和全长分子都可以在 PCR 中扩增。这些不完全 cDNA 的存在，极大降低甚至不能产生代表全长 cDNA 的 PCR 产物，因为在有异质和具相同末端底物的 PCR 反应液中，较小的 DNA 分子较易发生扩增。因此，应当优化逆转录过程，尽可能地避免不完全 cDNA 的产生，因而必须使用高质量的、未被降解的 mRNA。

在逆转录过程中降低 mRNA 二级结构稳定性的一个有效方法是在反应前提高反应混合物的温度（37～42℃）。嗜常温的逆转录酶（AMV2RT 和 MMLV2RT）只能在低于大约 50℃的温度下起作用。现已有热稳定的 DNA 聚合酶，可以在很高的温度（60～70℃）有活性。在此温度下，RNA 的二级结构不稳定，逆转录可以更容易进行到 mRNA 分子的 5′端。可以逆转录 mRNA 的嗜热 DNA 聚合酶主要有 *rTth*（PE 公司）和 *TetZ*（Amersham），有报道甚至 *Taq* 酶也表现出逆转录酶活性。虽然 *rTth* 酶较嗜常温的逆转录酶对二级结构不敏感，但也有不足。

标准的 RACE 要求把第一链 cDNA 用 TdT 酶加上一个脱氧核糖核酸同聚尾。然后，此同聚尾作为下步与锚多聚连接物杂交的底物。锚多聚连接物有一个 3′端同聚物锚（N）15，可与 cDNA 的同聚尾互补，然后由锚（N）15 引发第二链 cDNA 的合成。但此法也可能失败或促进无关产物的形成。为了使同聚加尾成功，逆转录后游离的核苷酸和过剩的 cDNA 引物必须除去。如果游离的核苷酸未除去，它将被 TdT 加入到 3′端 cDNA 尾中去，从而干

扰互补同聚物尾与锚引物的杂交。如果不除去过剩的 cDNA 引物，其 3′端将与 cDNA 末端竞争加尾，从而降低 cDNA 加尾的效率。加尾的 cDNA 引物还可以与加尾的锚引物退火，与 cDNA 发生竞争，从而降低第二链 cDNA 合成的效率。这些小分子将会在下步用锚引物 PCR 反应中作为退火和延伸的底物，降低 PCR 的效率。为了去除过剩的游离核苷酸和引物，可采用分子排阻类方法，如 Centricon microconcentrator（Amicon），但更简便、更经济有效的方法是使用 1/4 体积的 10mol/L 乙酸铵和 3 倍体积的 100％乙醇的两步沉淀法。5′-RACE 加尾的一个主要问题是所有的 cDNA，无论是不是全长，都被 TdT 加尾。一旦锚引物加到第二链合成中，这些切割后的 cDNA 片段会在 PCR 中与全长产物竞争扩增。如果这些 cDNA 比全长 cDNA 短得多的话，它们就可发生趋向性扩增，使得 PCR 反应不能积累足量的全长 cDNA，因而不能检测到目的产物，降低这些切断了的 cDNA 产生的唯一方法是优化逆转录过程。

（2）加尾的其他替代方法　加尾策略本身存在许多缺陷，有几个研究组独立地设计了 RACE 技术的改良方法以绕过 TdT 介导的 cDNA 加尾。所有这些改良方法都使用 RNA ligase 以共价方式把寡核苷酸锚连在 cDNA 的 5′端或直接连接在起始 RNA 群上。因此，这些策略既可消除 TdT 加尾这一步，又可消除 Anchor（N）15 引发的第二链 cDNA 的合成。这样，anchor（N）15 的内部引发可能性就得以完全消除。在把 anchor 直接连接在 RNA 上的方法中，从理论上讲，具有 3′端 anchor 序列的任何截断 cDNA 的产生都得以避免。有几个研究组发表了把 DNA 锚定寡核苷酸连接在 cDNA 上的方法。此种 RACE 变种通常称为锚连接 PCR（ligation-anchoredPCR，LA-PCR）或称单链 cDNA 末端的单链连接（single-strand ligation to ss-cDNA ends，SLIC）。

（3）PCR 扩增　因为 RACE 技术是为确定未知 cDNA 序列设计的，必须注意使在 PCR 中引入的突变最少。现已有具 3′→5′校对核酸外切酶活性的商品嗜热聚合酶可供使用，如 *Pfu*（Stratagene）和 *Vent*（New England Biolabs）等。

为了增加 PCR 反应的产率和特异性，已经设计了一些方法，如“热启动 PCR”，即先把不含模板或者模板和聚合酶的其他 PCR 组分混合，只有当混合物预热到引物退火温度以上时，再加入这些组分。

3. 注意事项

5′-RACE 的产物可能是单一产物、多个产物，甚至是不能分辨的连续条带，结果的质量取决于用来合成第一链和扩增的 GSP 的特异性、扩增所使用锚定引物的特异性、目的 mRNA 的复杂度和丰度及产物的长度。使用巢式引物扩增（最多可以三轮巢式扩增），并在连续几轮扩增中使用长度选择性的扩增产物作为模板可以增加 5′-RACE 的特异性。增加逆转录保温温度和 PCR 退火温度，并降低扩增反应中的镁离子浓度可以促进特异性。

5′-RACE 的灵敏度受所使用的逆转录酶和 cDNA 3′端抑制 cDNA 加尾的二级结构影响。cDNA 的不完全合成降低了全长产物的产量，并导致某些模板产生连续条带。因为 SuperScript Ⅱ 产生更多的全长 cDNA，所以可以提高 5′端序列检测的水平，尤其是对小于 1kb 的转录本。双链 3′端和发卡结构会通过减少用于加尾的 3′-羟基而破坏 cDNA 的加尾。cDNA 的起始保温温度在 94℃并随后在冰上冷却有助于破坏二级结构。一些困难模板可能需要加入 DMSO 辅助加尾，加尾酶末端脱氧核苷转移酶可以耐受最多 20％的 DMSO。

4. RACE 产物的克隆

一旦获得可以检测到的 RACE 产物，就必须进行纯化、回收和克隆，最好使用专门的

试剂盒回收。纯化的产物可用适当的限制酶切割，在低熔点琼脂糖胶上电泳，直接连接到测序载体中。现已有工程化的带有产生单个 T 碱基对突出的限制性位点的载体出售。在理想情况下，几个 RACE 克隆的测序就可表明目的 mRNA 的 5′端（或 3′端）的核苷酸序列。使用标准的 RACE 或 LA-PCR 克隆 5′端序列遇到的一个主要问题是截断了的 RACE 产物通常不能与全长 RACE 克隆相区别，建议使用 RLM-RACE 法，从至少两个 PCR 扩增产物中分离数个单一克隆以考察转录本的完整 5′端是否已被克隆。

三、基因组序列测定技术

DNA 测序一个反应最多只能测定 1000 个碱基，然后用已知序列的下游部分合成引物，进行另一反应，如此一步步地“步行”，逐步完成较大片段的测序。因此，需先用质粒建立许多克隆，构成质粒文库，再对这些质粒克隆进行测序，然后用电脑搭配成邻接克隆群。这里介绍两种大规模测序策略：基于 BAC 连续克隆系的测序和全基因组“鸟枪法”测序（shotgun sequencing）。前者是先把基因组打碎成 200～300kb 的片段并制成 BAC 文库，再选择一些 BAC 进一步打碎成 3kb 左右的小片段，测序并拼接。后者则绕过 BAC 克隆逐个排序的过程，把基因组 DNA 直接打碎成 3kb 左右的小片段，随机测序，然后利用这些小片段的重叠关系将其拼接成一条一致序列。

近年来测序技术发展很快，高通量的毛细管自动测序仪已经被广泛使用，且基于 BAC 连续克隆系的测序方法有向全基因组“鸟枪法”测序方法发展的趋势。下面，将详细介绍全基因组“鸟枪法”测序技术。

1. 测序主要步骤

先提取基因组 DNA、纯化（非常重要，要去除载体 DNA 或仅由载体 DNA 产生的片段），用限制性内切酶酶切、DNA 酶Ⅰ降解（加 Mn^{2+}）或超声波的方法将 DNA 打断、电泳，分别回收不同大小的 DNA 片段，构建质粒文库、转化宿主菌，扩增培养，提质粒作为模板，做 PCR 测序反应，上测序仪；处理所有测序数据，去除低质量和受载体污染的反应，拼接得到一组克隆群；利用正反向信息等确定这些克隆群之间的位置关系，补洞，最后得到一个完整的没有空隙（gap）的基因组一致序列。

2. 文库构建

首先将全基因组 DNA 打碎、电泳，回收 1.5～3kb 的片段，然后插入载体连接，构建质粒文库。将构建好的文库转化大肠杆菌（宿主菌），将大肠杆菌扩增培养即可得到大量可用于测序的模板 DNA。由于插入位点两侧序列相同，因此以两侧相同序列为引物对不同模板进行 PCR 测序反应，这种方法叫正反向末端测序法（pairwise end sequencing）。

3. 拼接

目前用于全基因组拼接的软件较多，例如 Phrap（http：//www. phrap. org）、CAP3、assembler、STROLL（http：//genetics. med. Harvard. edu/-tchen/STROLL）等。主要软件是由美国华盛顿大学 Phil Green 实验室开发的 Phred-Phrap-Consed 系统。Phred（测序器）是一种碱基识别系统（base-caller），它根据自动测序仪信号按顺序识别碱基，估计测序错误率等。Phrap（组装器）是根据 Phred 的结果从头组装由鸟枪法产生的不同的短序列。Consed（校对器）与 Phrep 组成一个有机整体，利用 Phrap 组装的序列由 Consed 编辑、整合人工校对结果等。

4. Scaffolding 和补空

当反应数达到一定覆盖率就可以结束随机测序阶段，当然由于建库质量不同，何时停止随机测序还应根据实际情况适时决定。然后，依照正反向反应的对应关系，确定克隆群的前后位置关系，这样一组前后位置关系确定的克隆群，就称为 scaffolds。构建 scaffolds 的过程即为 scaffolding。通过 scaffolding 后，所形成的 scaffolds 还不能拼成一个统一的完整基因组，它们之间还存在着洞，这些洞的位置关系并不确定，被称为 physical gap，而 scaffolds 内部的洞称为 sequence gap。

sequence gap 可以用横跨空隙的克隆 DNA 做模板，设计面向空隙的延伸引物，作 PCR 测序反应来填补；而最后剩下的 physical gap 则可在空隙两边克隆群末端设计出面向空隙的引物，然后将这些引物随机配对，以基因组总 DNA 为模板，进行 PCR 反应，以产物的有无来确定它们的位置关系。为保证基因组拼接的正确性，最后可以有选择地用几种限制性内切酶对基因组进行电子酶切，同时与实际的酶切图谱对照，如果两张图谱不符，则需重新拼接。所有测序工作完成后，应进行基因组注释工作，找出所有可能的基因，对它们的功能进行预测。

四、基因定位技术

目前国际上最为常用的基因定位方法是荧光原位杂交（FISH）和辐射杂种细胞系（RH）技术，下面将分别予以介绍。

（一）荧光原位杂交（FISH）技术

1. FISH 原理

荧光原位杂交（fluresence *in situ* hybridization，FISH）就是用标记有荧光分子的一小段 DNA 或 RNA 序列为探针，于合适的温度和盐离子浓度条件下，在原位与组织细胞内染色体上的互补序列特异性地退火结合而不破坏染色体的整体形态，以便在显微镜下观察到此特异的核酸序列及其在染色体上的位置的一种基因定位方法。FISH 技术分为 4 个主要步骤：第一步，将染色体、细胞或组织进行固定；第二步，标记探针；第三步，将探针与固定材料上的靶序列 DNA 或 RNA 杂交；第四步，检测杂交结果。

FISH 技术具有高度亲和力，有与放射性探针相同或更高的分辨率。现已可用不同的荧光染料同时进行多重原位杂交，显示出不同的荧光色泽。这种多色 FISH 技术近年来发展迅速，已成为基因定位作图和医学诊断的重要手段。染色体原位杂交的特点是杂交在显微镜载玻片染色体标本上进行。所谓原位即指标本上 DNA 原位变性，在利用非放射性标记的已知核酸探针杂交后，通过非放射性检测体系来检测染色体上特异 DNA 或 RNA 顺序，可用放射性颗粒在某条染色体的区带出现的最高频率或荧光的强弱来确定探针的位置，从而进行准确的基因定位。

2. 应用

FISH 提供了一种直接在细胞核中或染色体上确定 DNA 序列有无或相互间位置关系的方法。目前，可用 FISH 定位的 DNA 序列发展到了从 1kb 到整条染色体的范围；多色 FISH（multi-FISH）技术可在同一细胞核或中期染色体中显示不同的颜色，从而可同时检测两种以上的 DNA 三维结构。多色 FISH 制备的人类染色体光谱核型（spectral karotype），使得染色体自动化分析真正成为可能。用高度浓缩的中期染色体作 FISH 其分辨率很低（约 1～5Mb），而用拉长的间期染色质分辨率可提高到 50kb。由于对分辨率的要求不断提高，近年来又发展了许多新方法。如“自由染色质”（“free chromatin”）、纤维 FISH（fiber

FISH）和光学作图（optical mapping，OM）等，使得相距 1kb 的两个探针即可得以定位。在人类基因组计划中，FISH 起着分解染色体，使之易于测序，测序后再将 DNA 碱基序列有机地组织到染色体上的作用。

（二）辐射杂种细胞系（RH）技术

1. 原理

辐射杂种细胞系技术是一种体细胞杂交技术，是继荧光原位杂交技术后新近建立的染色体定位方法。RH 作图已成为当今构建人基因组大尺度、高精度、连续的染色体图谱的常用方法之一。该方法是 Cox 等在 Goss 等研究的基础上，利用高剂量的 X 射线将候选染色体打断成若干片段，这种染色体结构已被破坏的细胞可与仓鼠细胞形成杂交克隆。基于人染色体上的位标相距越近，其断裂频率越低的原理，通过 Southern 杂交和 PCR 的方法研究杂种克隆中人基因组 DNA 位标的分布，利用统计学的手段计算染色体位标断裂分离频率便可推算出位标之间的距离和在染色体上的排列，其精度范围为 0.3～30Mb。RH 技术自问世以来，在定位技术及统计学上均得以改进，从最初仅适于单条染色体作图，已逐渐发展成为整个基因组作图的一个有力手段。其实验操作程序为：①引物设计及合成；②PCR 反应；③数据的统计学分析。

2. 应用

FISH 法和 RH 技术是目前国际上最常用的基因定位的 2 种方法，各有其优缺点。FISH 可以定位基因组中多个同源的位点，结果直观、可靠，而 RH 法则很困难。但是 FISH 法检测步骤繁杂，尤其受探针大小的影响较大，2kb 以下的 cDNA 序列很难定位，而且结果需要十分有经验的细胞遗传学家进行分析。RH 法简单易行，定位精度高，适合于小至 1～2kb 以下的序列，根据统计学方法可直接进行染色体定位。

但是 RH 法也受到某些自身条件的限制，RH 嵌板中每个杂交克隆均包括仓鼠和人的基因组 DNA，可以看作是在仓鼠基因组 DNA“背景”上特异地扩增出人的 DNA 序列，因此引物序列应避免与仓鼠相似性过高，以免扩增出的条带无法与“背景”相区别。还应指出，RH 法能否成功定位依赖于该基因所在染色体区域上合适位标的存在。随着今后 RH 图谱精度的提高，越来越多的 STSs 被放置于染色体上，有望克服这个缺陷。

第二节　功能基因组学及常用研究技术

功能基因组学（functuional genomics）又往往被称为后基因组学（postgenomics），它利用结构基因组所提供的信息和产物，发展和应用新的实验手段，通过在基因组或系统水平上全面分析基因的功能，使得生物学研究从对单一基因或蛋白质的研究转向多个基因或蛋白质同时进行系统的研究。这是在基因组静态的碱基序列弄清楚之后转入对基因组动态的生物学功能学研究。研究内容包括基因功能发现、基因表达分析及突变检测。基因的功能包括：生物学功能，如作为蛋白质激酶对特异蛋白质进行磷酸化修饰；细胞学功能，如参与细胞间和细胞内信号传递途径；发育上功能，如参与形态建成等。采用的手段包括经典的减法杂交、差示筛选、cDNA 代表差异分析以及 mRNA 差异显示等，但这些技术不能对基因进行

全面系统的分析，新的技术应运而生，包括基因表达的系统分析（serial analysis of gene expression，SAGE）、cDNA 微阵列（cDNA microarray）、DNA 芯片（DNA chip）等。

鉴定基因功能最有效的方法是观察基因表达被阻断或增加后在细胞和整体水平所产生的表型变异，因此需要建立模式生物体（model organism）。比较基因组学（comparative genomics）是基于基因组图谱和测序基础上，对已知的基因和基因组结构进行比较，来了解基因的功能、表达机理和物种进化的学科；利用模式生物基因组与人类基因组之间编码顺序上和结构上的同源性，可以克隆人类疾病基因，揭示基因功能和疾病分子机制，阐明物种进化关系，以及基因组的内在结构。

一、基因突变检测（分析）技术

基因突变的研究已成为当今生命科学研究的热点之一，检测方法也随之迅速发展。基因突变检测可以从两个层次来入手：一是蛋白质层次的间接检测，主要是借助一些蛋白质分析技术，分析突变体产生的变异蛋白，以及由变异蛋白引起的其他改变，比如从菌落形态、抗生素抗性等来入手；二是可以对遗传物质（主要是 DNA）的直接检测，一般来说，这类技术是通过建立一系列电泳技术，分析 DNA 构象或解链特性，或者利用 DNA 变性和复性等特性，来进行 DNA 突变的分析。目前由于 PCR 技术的应用，基因突变检测主要以分析 DNA 变异的直接检测为主。

（一）单链构象多态性（SSCP）技术

单链构象多态性（single-strand conformation polymorphism，SSCP）技术是一种较早应用的基因突变检测方法。该方法对检测基因的单个碱基置换和某一片段 DNA 突变位点的筛查提供了有效而快速的手段。

1. SSCP 的原理及特点

日本 Orita 等研究发现，单链 DNA 片段呈复杂的空间折叠构象，这种立体结构主要是由其内部碱基配对等分子内相互作用力来维持的，当有一个碱基发生改变时，会或多或少地影响其空间构象，使构象发生改变，空间构象有差异的单链 DNA 分子在聚丙烯酰胺凝胶中受排阻大小不同。因此，通过非变性聚丙烯酰胺凝胶电泳（PAGE），可以非常敏锐地将构象上有差异的分子分离开，作者称该方法为单链构象多态性分析。在随后的研究中，作者又将 SSCP 用于检查 PCR 扩增产物的基因突变，从而建立了 PCR-SSCP 技术，进一步提高了检测突变方法的简便性和灵敏性。其基本过程是：①PCR 扩增靶 DNA；②将特异的 PCR 扩增产物变性，而后快速复性，使之成为具有一定空间结构的单链 DNA 分子；③将适量的单链 DNA 进行非变性聚丙烯酰胺凝胶电泳；④最后通过放射性自显影、银染或溴化乙锭显色分析结果。若发现单链 DNA 带迁移率与正常对照的相比发生改变，就可以判定该链构象发生改变，进而推断该 DNA 片段中有碱基突变。

2. SSCP 的改进

刚建立时是将同位素掺入 PCR 扩增物中，通过放射自显影来显示结果。这给该技术的推广造成一定的困难，随着 DNA 银染方法与 PCR-SSCP 结合，尤其是直接溴乙锭染色方法的应用，使得该方法大大简化。最近，SSCP 值得注意的改进是将 DNA-SSCP 分析改为 RNA-SSCP 分析，其基本原理是：RNA 有着更多精细的二级和三级构象，这些构象对单个碱基的突变很敏感，从而提高了检出率。另外，RNA 不易结合成双链，因此可以较大量的进行电泳，有利于用溴化乙锭染色。但该方法增加了一个逆转录过程；还需要一个较长的引

物，内含有启动RNA聚合酶的启动序列，从而相对地增加了该方法的难度。为了进一步提高SSCP的检出率，可将SSCP分析与其他突变检测方法相结合。其中，与杂交双链分析（heterocluplex analysis，Het）法结合可以大大提高检出率。Het法是用探针与要检测的单链DNA或RNA进行杂交，含有一对碱基对错配的杂交链可以和完全互补的杂交链在非变性PAGE上通过电泳被分离开。为了高灵敏特异性地显示SSCP分析结果，现已发展多种PCR-SSCP技术，各有其优势及适用领域；例如，可以在PCR扩增中用同位素或荧光素等标记引物或核苷，也可以在电泳后用银染或溴化乙锭染色以显示结果。较为常用的放射性同位素PCR-SSCP法。在进行PCR扩增特定靶基因序列时，利用γ-^{32}P-ATP标记引物或直接在PCR反应体系中加入α-^{32}P-dCTP进行PCR扩增，使扩增产物带有同位素标记物，然后将扩增产物变性为单链进行非变性聚丙烯酰胺凝胶电泳，放射自显影显示结果。利用引物标记或碱基掺入法使PCR扩增产物带有同位素标记物，均可使产物信号增强几个数量级。二者相比较，前者经济、扩增特异性强，多用于大样本的检测和筛选；后者操作简便，适于一般实验室开展，可用于小样本或大样本的检测和筛查。

3. 要点及注意事项

（1）核酸片段的大小　用于SSCP分析的核酸片段越小，检测的敏感性越高。对于<200bp的片段，SSCP可发现其中70%的变异；对于300bp左右的片段则只能发现其中50%的变异；而>500bp的片段，则仅能检出10%～30%的变异，因此，<300bp，尤其是150bp左右的核酸片段更适于SSCP分析。对于大于400bp的PCR产物就需要设法进一步处理，可以用限制酶消化PCR产物，产生小于400bp的DNA片段，再进行SSCP分析。

（2）游离引物　游离引物可能同PCR产物结合而改变其泳动率，即使游离引物量为6nmol/L都有明显影响。因此，应尽可能除去游离引物。可以采用不对称引物扩增方法，尽可能消耗多余的引物；也可以运用过柱或磁性球方法纯化PCR产物；或者是稀释PCR产物，减少游离引物的干扰。

（3）低浓度变性剂　凝胶中加入低浓度的变性剂，如5%～10%甘油、5%尿素或甲酰胺、10%二甲亚砜（DMSO）或蔗糖等有助于提高敏感性，可能是因为轻微改变单链DNA的构象，增加分子的表面积，降低单链DNA的泳动率。但有些变异序列却只能在没有甘油的凝胶中被检出。因此，对同一序列使用2～3种条件做SSCP，可能提高敏感性。

（4）电泳温度　一般认为保持凝胶内温度恒定是SSCP分析最关键的因素，温度有可能直接影响DNA分子内部稳定力的形成及其所决定的单链构象，从而影响突变的检出。室温下电泳适于大多数情况，但由于在电泳时温度会升高，为确保电泳温度相对恒定，应采取以下措施：减少凝胶厚度，降低电压，有效的空气冷却或循环水冷却等。

（5）凝胶的长度　可用测序板进行SSCP分析，凝胶板长度在40cm以上。

（6）凝胶浓度及厚度　凝胶浓度很重要，一般使用5%～8%的凝胶，凝胶浓度不同，突变带的相对位置也不相同，如果在进行未知突变种类的SSCP分析时，最好采用两种以上凝胶浓度，这样可以提高突变种类的检出率。凝胶的厚度对SSCP分析也很重要，凝胶越厚，背景越深，在上样量较多的前提下，尽量使凝胶越薄越好。

（7）假阴性　一般认为，如没有污染，PCR-SSCP分析不存在假阳性结果，但可能出现假阴性结果，后者是由于点突变引起的空间构象变化甚微，迁移率相差无几所致，尤其是点突变发生在扩增片段的两端时。如果有阳性和阴性对照，结果可以重复确定的突变带是可信的，如果没有阳性对照，应经测序来确定其是否为突变带。由于PCR-SSCP的不足之处主

要是可能检出假阴性结果。应通过设置阳性对照，摸索电泳条件，假阴性结果在很大程度上是可以避免的。但对未知基因变异的检测，假阴性结果就难以百分之百地消除。

(8) 结果分析 单链凝胶电泳时，互补单链迁移率不同，一般形成两条单链带。PCR产物进行单链凝胶电泳之前，通过加热变性产生单链。如变性不彻底，残留双链亦可形成一条带。因此，PCR-SSCP分析结果至少显示三条带。但是，由于一种DNA单链有时可形成两种或多种构象，检出三条或四条单链带也就不足为奇。

4. SSCP的应用

自从Orita等用SSCP进行人类DNA多态性分析以来，该方法大量地用于检测和肿瘤发生有关的基因突变，例如检测星形细胞瘤、脑瘤、小细胞肺癌、胃癌、肠癌等肿瘤中的*p53*基因的突变、肺癌的*ras*基因突变等。最近Sugano等用银染色SSCP法，成功地检测了c-Ki-*ras* 2基因第12位的突变，电泳和银染色过程在2.5h内完成。SSCP还用于检测引起人类遗传性疾病的研究，如在囊性纤维化中起作用的CFTR基因、神经纤维瘤1型基因、家族性结肠息肉基因的研究中。最近，Sharkar等用RNA-SSCP法检测了28名B型血友病患者的凝血因子Ⅸ的基因序列，并与DNA-SSCP和直接基因测序法进行了比较，发现在全长2.6kb的凝血因子Ⅸ基因中，有20处碱基点突变，RNA-SSCP可检测出其中的70%，而DNA-SSCP只能检测出35%，显示了RNA-SSCP比DNA-SSCP有更高的灵敏性。

SSCP法除大量地用于基因突变的检测外，还用于病毒的分型、监测PCR实验中的污染情况以及病原体传播途径的研究中。Yap用巢式PCR对来自不同国家和地区的HBV标品进行了检测，并对扩增产物进行了SSCP分析，发现每个标本的SSCP图谱完全不同，不仅证明了HBV DNA具有地区性变异，而且排除了交叉性污染的可能性，为保证PCR诊断结果的可靠性找到了好方法。另外，Yap等还将SSCP用于DNA定量分析上，他们将SSCP与竞争性PCR法相结合进行乳腺癌细胞突变*p53*基因的定量分析，并取得了初步成功。该方法的优点在于内标和待测DNA只有一个碱基不同，这样使内标和待测DNA有相同的扩增条件，从而更准确地测出DNA的量。从以上可以看出，SSCP正在分子生物学领域发挥着巨大的作用。

(二) 变性梯度凝胶电泳 (DGGE) 技术

变性梯度凝胶电泳 (denaturing gradient gel electrophoresis, DGGE) 技术是检测基因突变较为精确的方法，它不仅可检测单一片段的单点突变，对多点突变的检测亦较容易，该方法与PCR技术相结合，使其能非常快速地对大量标本进行分析。该方法经改进，又发展出了恒定变性凝胶电泳 (constant denaturant gel electrophoresis, CDGE)、瞬时温度梯度电泳 (temporal temperature gradient electrophoresis, TTGE) 和温度梯度凝胶电泳 (temperature gradient gel electrophoresis, TGGE)。

1. 原理

DGGE主要是利用梯度变性胶来分离DNA片段。电泳开始时，DNA在胶中的迁移速率仅与分子大小有关，而一旦DNA泳动到某一点时，即到达该DNA变性浓度位置时，使得DNA双链开始分开，从而大大降低了迁移速率。由于不同的DNA片段的碱基组成有差异，使得其变性条件产生差异，从而在凝胶上形成不同的条带。目前常用的变性剂有尿素 (urea) 和甲酰胺 (formamide)。

根据DGGE变性梯度方向与电泳方向是否一致，可将其分为两种形式：垂直DGGE和平行DGGE。垂直DGGE的变性梯度方向与电泳方向垂直，可用于优化样本的分离条件，也可用于分析PCR产物的组成；平行DGGE的变性梯度方向与电泳方向一致，可用于同时

分析多个样本。检测某种突变的变性剂浓度一般通过变性梯度凝胶电泳来确定。

(1) 恒定变性凝胶电泳 (CDGE) 恒定变性凝胶电泳技术是在 DGGE 基础上发展起来的基因突变检测技术，其主要化学物质和基本方法都一样，亦是利用突变基因与正常基因片段间 T_m 值的差异来检测突变的方法，该方法与 DGGE 的区别在于聚丙烯酰胺中含的变性剂浓度相当于含突变基因片段的 T_m 值，而且浓度恒定，而不是线性递增。这使得整个电泳变得更加简单快速。为了提高 DGGE 和 CDGE 的检测敏感性，通常在一侧引物的 5′端设计一含 GC 的区域 (GC-clamp，GC 钳)，避免了 DNA 的完全解链，从而提高了突变的检出率，可达 100%。

(2) 瞬时温度梯度电泳 (TTGE) 其检测突变的能力和 DGGE 差不多，但却不需要变性梯度胶。基本方法是：在加有一定浓度尿素的聚丙烯酰胺凝胶电泳过程中，使外界温度恒定地增加，这样在跑胶的过程中形成一个温度梯度，变性环境是由恒定浓度的尿素和瞬时温度梯度共同形成。这样使得全过程更加简单迅速，分辨率极高。

(3) 温度梯度凝胶电泳 (TGGE) 温度梯度凝胶电泳技术是将 PCR 扩增产物加入中性聚丙烯酰胺凝胶上，在一温度线性梯度上升的电场中进行电泳，当 DNA 双链到达其 T_m 时，双链解开，形成分叉，而影响电泳迁移率，突变的扩增产物其 T_m 值与野生型有明显不同，因而有不同的解链区，从而造成电泳迁移效率的差别，据此可判断检材中 DNA 有无突变。其基本原理实际上与 DGGE 差不多，只是由变性剂形成的梯度被温度梯度所代替。这样的梯度可由微处理器控制，与 DGGE 相比，更加稳定可靠。如在变性高压液相色谱 (denaturing high pressure liquid chromatography，DHPLC) 中，其杂合双链和纯合双链的分离就是通过精确的温度控制，使杂合双链部分变性，从而与纯合双链分离开来。

2. 基本操作

DGGE、CDGE、TTGE 和 TGGE 的操作基本一样，美国的 Bio-Rad 公司开发出的 Dcode 普通突变检测系统将前三者与 SSCP 合而为一，给使用者带来了极大的方便。Dcode 系统有一个温度调节装置，该装置由微处理器控制的加热器、缓冲液循环泵和搅拌器所组成，温控范围为 5～70℃，如果电泳温度要超过此范围，只要外接一个制冷器即可，十分方便。该温度调节装置的另一个好处是十分适应实验室的需要，如在 65℃进行了 DGGE 分析，现在要换成 10℃下的 SSCP，只要在系统中加入 SSCP 的成套装置即可。其他配置和普通电泳的装置差不多；操作上和 SSCP 相似，但对条件的控制更容易，操作更方便；固定、染色、显色程序与普通电泳一样。

3. 主要优缺点

优点：①突变检出率高。在 TGGE 中，一般 2mm 可对应 0.6℃温差，而在 TTGE 中，每小时升温可控制为 0.1℃，极小的 T_m 值的差异也可以检测出来，尤其适用于单对碱基突变个体的筛查；②检测片段可达 1kb，尤其适于 100～500bp 的片段 (而 PCR-SSCP 一般只适用于 150～200bp，对较大片段的突变检出率将大大降低)；③加样量小，仅需 1～5ng 的 DNA；④操作简便，快速；⑤重复性好，因为主要电泳条件都由微处理器控制。

该法的缺点是：①DGGE 对肿瘤的突变检测有困难，尤其是实体瘤；②一般不能确定突变位置，最终还得依赖于 DNA 测序。

4. 在基因组研究中的应用

对未知基因突变的筛选 PCR-DGGE 可以准确鉴定单碱基替换、移框突变和少于 10 个碱基的缺失突变。一般来讲，就是用病人的基因组 DNA 或者是有突变个体的 DNA 进行 PCR

扩增，然后用 DGGE 分析。为提高检出率，常将突变的 DNA 和正常人的 DNA 混合变性，这样电泳后可形成 4 条带。该法尤其适用于大样本的筛查，对一些罕见突变型的筛查十分有用。利用 Riesner 等设计的 PCR-TGGE，根据条带位置的改变，利用热动力学统计方程还可计算确定突变发生的位置。

(1) 在突变谱和拷贝数确定中的应用　DGGE 的另一个作用是可分析突变的组成，即突变谱 (mutational spectra)，主要是利用诱变剂体外诱变人类原始淋巴细胞株，使之生成大量的 $HPRT^-$ 突变体，然后进行 PCR 扩增，用 DGGE 分析。用定量 PCR-TGGE 方法还可以确定标本中某特异 DNA 序列的拷贝数。其原理是将已知拷贝数的内参照与标本中的靶基因共同扩增，然后 TGGE 分离两者扩增产物，根据其比例可以确定标本中靶基因的拷贝数。

(2) 在蛋白构象研究中的应用　TGGE 还可用于蛋白质分子结构和热稳定性研究，对于某些温敏感蛋白的分离和鉴定也十分有用。此外在临床基因诊断中，虽然通过突变筛选可以确定基因与疾病的关系，但仍然需要通过突变分析来寻找疾病基因在不同患者中的突变特点或热点，以利于临床基因诊断和产前诊断，而对于疾病基因较大、不存在突变特点或热点的遗传病，除了应用连锁分析法进行基因诊断和产前诊断之外，本文所介绍的方法是进行突变筛查和产前诊断的又一重要选择。

(三) 直接测序法

通过对同个体同一基因或片段进行测序和序列比较，寻找所研究的序列中存在的所有碱基变异，其检出率可达 100 %。DNA 序列分析是检测基因突变最直接最可信的方法，它不仅可确定突变的部位，而且还可确定突变的性质。现在的测序方法已与经典的方法有了很大的不同，其基本原理虽仍是双脱氧终止法，但自动化程度大为提高，操作更简便，测序时间也大大缩短。随着 PCR 技术与测序联合使用，不需经过亚克隆步骤，故称为直接测序法 (direct sequencing，DS)。PCR 直接测序是指对 PCR 产物进行的直接序列分析，而不是传统的测序技术（先将待测 DNA 片段克隆于测序载体上），这不仅大大简化了操作步骤，节省大量的人力和物力，而且可实现自动化操用，加之新的荧光检测技术的应用，使测序的效率大大提高。直接测序一般对两种纯合子进行，但对杂合子进行了直接测序也可达到目的。随着测序的自动化程度的提高和测序成本的降低，直接测序将会越来越多地用于“SNP”的检测与分型。但目前直接测序仍然为一种费用较高的方法。

采用 PCR 循环直接测序时，应首先将扩增产物转化为单链测序模板，目前常用的转化方法为不对称 PCR，即在反应体系中通过引物浓度的差异来形成单链 DNA，通常两侧引物的浓度为 100∶1。当一侧引物被耗尽后，另一侧引物扩增的片段即为单链，然后即可用于测序；此外，获得单链 DNA 还有磁珠俘获法、外切酶消化法及 genomic amplification with transcript sequeucing 法（简称 GAWTS 法）。PCR 直接测序的最新发展是将双脱氧终止子技术与 PCR 技术相结合进行循环测序，在 PCR 反应体系中同时将 ddNTP 加入，并利用同位素或荧光素标记的引物引导扩增后，使模板的扩增与测序同时进行，其特异性在于使用的模板量少且不需分离单链。

应用 PCR 测序有以下优点：①模板需要量少；②方法简便，操作易于标准化、自动化；③测序效率高、准确，在短时间即可完成。

测序对突变的检测准确，但耗费人力、物力较多。随着人类基因组计划的不断深入，很多高效、快速的测序方法被开发应用。在大规模突变检测中，多重 PCR 测序和利用基因芯

片测序有良好的应用前景。

（四）单碱基延伸标签阵列技术（SBE-TAGS）

SBE-TAGS 是以具有突变位点的 PCR“产物作为模板，加入 SBE”引物及用不同荧光标记 ddNTPs 进行延伸反应；每个引物的 5′端连接一个特异标签，而其 3′端正好终止于突变位点的前一个碱基处，这样延伸出的一个碱基就是突变碱基。将 SBE 延伸产物与固相支持物上和特异标签互补的探针阵列作用（见图 11-2），由于 4 种 ddNTPs 分别标记不同的荧光染料，反应结束后通过分析荧光的颜色即可确定突变的为何种碱基。通过对许多不同的 SBE 引物连接不同的特异标签可一次性对大量目的片段进行检测，并且可用玻璃板等材料做杂交阵列的固相支持物，制作费用较低。其不足在于特异标签的选择比较复杂，工作量较大；同时在结果的处理上也不够简便直接。

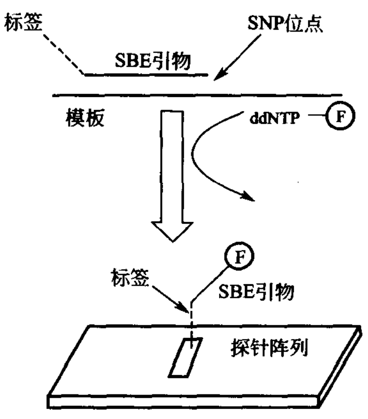


图 11-2 SBE-TAGS 技术原理示意图

（五）异源双链分析技术

异源双链分析（heteroduplex analysis，HA）的原理为：在同一扩增体系中，若同时含有正常模板和突变型模板时，随着 PCR 的循环进行，野生型基因片段和突变型片段均可扩增，并形成异源双链，由于异源双链中错配区的出现，因而其构型亦发生改变，这种构型上的差别会形成电泳图谱上的差异，即由于突变和野生型 DNA 形成的异源杂合双链 DNA 在其错配处会形成一突起，在非变性凝胶中电泳时，会产生与相应的同源双链 DNA 不同的迁移率。该法与 SSCP 相似，所不同的是 SSCP 分离的是单链 DNA，HA 法分离的是双链 DNA，也只适合于小片段的分析，一般在 300bp 以下，其敏感性与 SSCP 相似。该技术简便，适用于快速的突变检出，已用于多种遗传病基因突变的研究。但 HA 对一些不能用 SSCP 检出的突变有互补作用，两者结合使用，可使突变检出率提高到近 100%。

（六）连接酶链反应

与其他核酸扩增技术比较，连接酶链反应（ligase chain reaction，LCR）的最大特点为可准确区分基因序列中单个基因突变，由 Landegree 于 1988 年首次应用于镰刀型细胞贫血的分子诊断。LCR 是以 DNA 连接酶将某一 DNA 链的 5′-磷酸与另一相邻链 3′-羟基连接为基础，应用两对互补的引物，双链 DNA 经加热变性后，两对引物分别与模板复性，若完全互补，则在连接酶的作用下，使相邻两引物的 5′-磷酸与 3′-羟基形成磷酸二酯键而连接，前一次的连接产物又作为下一次循环反应的模板，如果配对的碱基存在突变则不能连接和扩增。LCR 产物检测最初是通过这 ^{32}P 标记上游引物 3′末端，经变性凝胶电泳分离后放射自显影加以鉴定，其检测敏感性达到 200 个靶分子。也可设计 1 个横跨两引物的检测探针，用它与 LCR 产物进行杂交检测。近年有应用荧光素、地高辛等非核素标记方法。Batt 在 1994 年发展了一种更为简的方法，称微孔板夹心杂交法。由于 LCR 的快速、特异和敏感的特性，以及能检测单个碱基突变的能力，因此被应用于肿瘤基因突变的分子诊断，并与 PCR 结合用以提高其敏感性。

（七）等位基因特异性寡核苷酸杂交（ASOH）

如果一个基因的突变部位、性质经测序分析已经阐明，即可用该方法直接检测突变。该

方法的原理即利用人工合成的寡核苷酸片段（一般为 19base）作为探针，与经 PCR 扩增获得的靶 DNA 进行杂交，在严格控制杂交条件的前提下，通过斑点杂交或其他类型的杂交来检测 PCR 产物中有无无相应突变，即探针与靶 DNA 片段之间只要有一个碱基不相互配对或错配，就能检出。

（1）探针　探针一般合成两个，一个为正常序列，另一个含有突变位点，前者与正常靶 DNA 完全互补，后者只与突变 DNA 互补，一般长度为 19base。该探针称之为等位特异寡核苷酸探针（allele specific oligo-nucleotide probe，ASO 探针）。

（2）杂交类型　PCR 技术的出现从根本上解决了靶 DNA 的来源问题，使人们能在很短的时间内获足够量的靶 DNA，传统的 PCR-ASO 技术主要是将 PCR 产物固定于杂交膜上，然后用不同的标记探针检测，亦有将已标记好的探针预先固定于杂交膜上，再使之与 PCR 产物结合，检测有无完全互补的 DNA 产物。

（3）探针的标记　最常应用的标记物为放射性物质有 ^{32}P、^{34}S，但由于其来源、半衰期及放射性危害不能普遍应用。PCR 技术的引进使可在短时间内获得足够量的靶 DNA 片段，因而对探针的要求亦有所降低；非同位素标记的探针亦可获非常满意的结果，它不仅使 ASO 探针使用起来更快更简便，且无同位素半衰期的限制，操作亦很安全，适合于临床应用。

（八）RNA 酶 A 切割法

在一定条件下，异源双链核酸分子 RNA∶RNA 或 RNA∶DNA 中的错配碱基可被 RNase A 切割（RNase A cleavage），切割产物可通过变性凝胶电泳分离。当 RNA 探针上错配的碱基为嘌呤时，RNaseA 在错配处的切割效率很低，甚至不切割；而当错配碱基为嘧啶时，则其切割效率较高。故如果仅分析被检 DNA 的一条链，突变检出率只有 30%，如同时分析正义和反义两条链，检出率可达 70%。该法需要制备 RNA 探针，增加了操作的复杂性，但可用于 1～2kb 的大片段进行检测，并能确定突变位点。由于这些优越性，它仍被作为一种经典方法用于对未知突变进行分析。

（九）基于 PCR 与酶切的技术

最常用的是 PCR-RFLP（限制性片段长度多态性）和随机扩增多态性 DNA（random amplification polymorphism DNA，RAPD），前者通过专一识别性内切酶处理 PCR 产物，电泳后检测分型，后者利用随机引物（约 10base）对基因组进行扩增，寻找特征性条带和分型；此外，还有突变酶学检测（enzymatic mutation detection，EMD）和扩增阻滞突变系统（amplification refractory mutation system，ARMS）等。

（1）突变酶学检测（EMD）　EMD 是利用 T4 内切酶Ⅶ的特性检测 SNP，T4 内切核酸酶是一类能识别结构而不是序列的酶，它能够识别并切割一些特殊 DNA 底物，如十字结构、分支 DNA、单链突出、单碱基错配而形成的泡状结构及因此而形成的泡突。经 T4 内切酶Ⅶ处理后的产物片段可经电泳分离，根据片段大小判断有无 SNP 的存在，EMD 可分析几个 kb 的 DNA 片段，且速度快、方便、检出率高。

（2）扩增阻滞突变系统　扩增阻滞突变系统又称为等位基因特异扩增（ASA）或等位基因特异性 PCR（AS-PCR）。其原理是：由于 *Taq* 酶缺乏 3′→5′外切酶活性，所以如果引物 3′端不能与模板严格互补配对时，模板 DNA 就不能有效扩增。一般的 ARMS 需两对 PCR 引物，其中一对为正常引物（NP），另一对为突变引物（MP），且 MP 和 NP 分别用不同的荧光进行标记，同时加入反应体系，两类引物与模板竞争结合，PCR 结束后纯合的野

生型样品只能够显现正常引物的颜色，纯合的突变性样品只能够显现突变引物的颜色，而杂合样品可以与两对引物结合，因而两种颜色都有，且几乎等量。

ARMS 于 1989 年建立，主要用于对已知突变基因进行检测。该法通过设计两个 5′端引物，一个与正常 DNA 互补，一个与突变 DNA 互补，对于纯合性突变，分别加入这两种引物及 3′端引物进行两个平行 PCR，如有与突变 DNA 完全互补的引物才可延伸并得到 PCR 扩增产物。如果错配位于引物的 3′端则导致 PCR 不能延伸。ARMS 借鉴多重 PCR 原理，可在同一系统中同时检测两种或多种等位基因突变位点。其检出率依赖于反应条件的优化和可能发生的引物与靶 DNA 的错配延伸，特别是当错配碱基为 G：T 时，这时可通过调整实验条件，如引物靶 DNA、*Taq* DNA 聚合酶的浓度等来获得高特异性。在反应体系中加入甲酰胺也可减少非特异性扩增。还可通过在引物 3′端的第二个碱基，引入一个错配碱基，使之与模板之间形成双重错配以阻止错误延伸。

（十）高通量检测技术

1. 基因芯片技术

基因芯片（gene chip）又称 DNA 芯片（DNA chip）、DNA 微阵列（DNA microarray）等，指采用寡核苷酸原位合成或显微打印手段，将大量的 DNA 片段有序地固定排列在固相支持物表面形成探针阵列，然后与标记的样品进行杂交，对杂交信号的检测实现了快速、高效、并行的多态信息分析。David G. Wang 等运用差异检测阵列（variant detector array，VDA）检测并发表了第一张人类基因组单核苷酸多态性（SNPs）遗传连锁图，VDA 对于一个目的片段设计包含 4 种与之互补的探针阵列，每种探针仅在与候选 SNPs 位点配对的一个碱基不同（分别为 A、G、C、T 中的一种），然后将蛋白毒素标记的 PCR 扩增产物与阵列上的探针进行杂交，用计算机收集并分析荧光信号即可得到结果；大量 VDA 集成在一个芯片上，一次可同时检测 500 多个候选片段。在基因突变检测方面 DNA 芯片也有广阔的前景，其基本原理为将许多已知序列的寡核苷酸 DNA 排列在 1 块集成电路板上，彼此之间重叠 1 个碱基，并覆盖全部所需检测的基因，将荧光标记的正常 DNA 和突变 DNA 分别与 2 块 DNA 芯片杂交，由于至少存在 1 个碱基的差异，正常和突变的 DNA 将会得到不同的杂交图谱，经过共聚焦显微镜分别检测两种 DNA 分子产生的荧光信号，即可确定是否存在突变，该方法快速简单、自动化程度高，具有很大的发展潜力，将在基因突变检测中发挥非常重要的作用。基因芯片具有信息量大和自动化程度高的突出优点。但它也存在若干问题，例如芯片造价高昂，所需设备贵重，不利于普及应用，一次要进行大量样本的检测，相应的准备工作量也大大增加。

2. TaqMan 技术

TaqMan 技术运用了荧光共振能量转移的原理，即供者-受者（引爆者-猝灭者）染料对彼此靠近时供者所发荧光由于光谱接近受者而被猝灭，但两者分离时则可检测到供者所发荧光。设计能与 PCR 产物杂交的探针，其 5′与 3′端分别标记供者-受者染料对，正常情况下检测不到供者分子发出的荧光，但当溶液中有 PCR 产物时，探针与之相结合，在进行延伸反应时激发 TaqMan DNA 聚合酶的 3′外切酶活性，将探针 5′端连接的荧光分子切割下来而发荧光。如果探针与目的序列存在错配碱基，则二者结合的紧密程度及酶切割供者的活性都会降低，相应地也影响了释放荧光的强度，从而可通过检测反应液中荧光的强度来确定 SNPs。TaqMan 将检测结合在 PCR 反应过程中进行，无需分离或洗脱过程，提高了速度，并减少了 PCR 污染的可能性；但不能采用 2 个及 2 个以上的探针同时进行检测；所采用荧

光退火及双末端标记技术，退火难以彻底。

3. 双色荧光偏振检测入侵技术（IADCFP）

设计入侵探针、初始探针及信号探针与目的序列的 PCR 扩增产物共孵育。入侵探针与目的片段突变位点的下游序列互补，其 3′端正好终止于突变位点；初始探针则与突变位点的上游序列互补，如果它在突变位点也与目的序列互补，则会与入侵探针产生重叠，这一结构能被特定的裂解酶识别并在该处切断初始探针，被切下的片段入侵具有发夹结构的信号探针，裂解酶将识别信号探针上形成的重叠结构并切掉其 5′端带偏振荧光标记的一段序列（见图 11-3）。荧光标记被平面偏振光激活后发出偏振信号，由于任何明显的分子量改变都会改变信号的偏振状况，可以通过对荧光偏振状况的分析很容易地识别目的序列的突变。

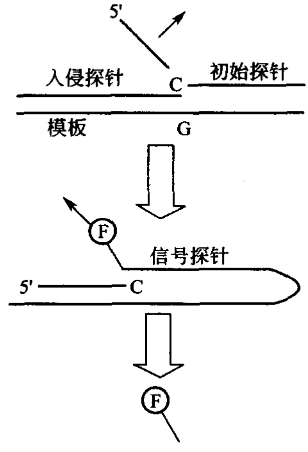


图 11-3 IADCFP 技术原理示意图

IADCFP 简便易操作，不需要 PCR 产物的提纯和凝胶电泳，并且可在一个反应中完成检测，结果非常可靠，但是对每个 SNP 的 2 个等位基因需要 2 个不同的实验，增加了时间的消耗和工作强度。

二、比较基因组杂交技术

比较基因组杂交（comparative genomic hybridization，CGH）技术是在染色体荧光原位杂交（FISH）技术的基础上发展起来的一种新的分子细胞遗传学技术，是检测细胞染色体非平衡性异常的一种新方法。

CGH 的基本原理是：分别用不同的荧光标记体系标记来自待检组织和正常组织的全基因组 DNA（分别称为测试 DNA 和参照 DNA），各取等量标记产物制备成混合探针，与足够量的同一种属来源的 Cot-1 DNA 先进行预杂交以封闭基因组上的分散重复序列（interspersed repetitive sequence，IRS），然后与正常中期分裂相进行染色体原位抑制杂交。测试 DNA 探针和参照 DNA 探针竞争性地与染色体上的靶序列杂交，经计算机软件分析，将染色体上每一像素上测试 DNA/参照 DNA 荧光强度的差异进行换算，以研究测试 DNA 拷贝数的增多或缺失。该技术不需细胞培养，一次杂交实验即可在整条染色体或染色体区带水平对不同基因组间 DNA 序列拷贝数的差异进行检测并定位。

1. 比较基因组杂交方法要点

（1）正常分裂中期细胞的制备　从正常人中获得的周围血淋巴细胞加入植物血凝素进行培养，用秋水仙碱处理使其停留于有丝分裂中期。用低渗氯化钾溶液和甲醇与醋酸的混合液处理后，将细胞置于玻片上即可获得分裂中期细胞的染色体。

（2）DNA 提取　应用常规的 DNA 提取方法，从大约 2×10^6 个细胞或 $1mm^3$ 的组织中提取高分子量 DNA。从待测细胞中提取高质量 DNA 至关重要。如检测肿瘤，所取标本需具代表性，所含肿瘤细胞比例要高，且尽可能不含污染的坏死物质。制备高分子量 DNA 可用蛋白酶 K-酚提取、乙醇沉淀；福尔马林固定、石蜡包埋的病理切片可用显微切割法取得肿瘤细胞，提取 DNA，再进行变性引物介导的聚合酶链反应（DOP-PCR）扩增靶 DNA。从石蜡包埋的组织中提取 DNA 可能会遇到 DNA 降解的问题，但可以通过后面的标记方法加以弥补。用于对照的正常 DNA 可以从健康人周围血淋巴细胞中获得。

(3) DNA 探针标记　目前最常用的方法是酶介导的切口平移法。DNA 探针可间接标记(以生物素-14-dUTP 和地高辛-11-dUTP 标记，用免疫组化法检测)，也可直接标记［用连接有荧光素的分子如罗丹明-dUTP 和异硫氰酸酯 (F ITC)-dUTP 分别标记待测和参照DNA，直接检测杂交］。与间接标记法相比，直接标记的信号较弱，但其信号的颗粒性较低，减少了杂交背景的干扰，且直接荧光标记可解决 1p 和 19 号染色体假阳性丢失的问题。标记前后 DNA 片段的大小是整个实验的关键。为获得最佳的标记效果，DNA 应控制在 500～2000bp 之间。对一些大分子量的 DNA 在标记前可以先予机械剪切。

(4) 原位杂交　标记好的待测、参照 DNA 探针各 100～200μg 与 15～20mg·的人 Cot-1DNA (起封闭 *Alu*、*Kpn* 等重复序列以降低本底的作用) 相混合，沉淀后用含 50%甲酰胺和 10%右旋糖酐的杂交液 10μl 溶解。分别将探针和分裂中期细胞的染色体玻片进行热变性处理。在 37℃，湿性的恒温箱杂交 2～4d，严格冲洗，再以 4,6-二氨基-2-苯吲哚 (DAPI) 复染，如是间接标记探针，则用连有不同荧光的抗体与不同标记的半抗原通过免疫组化染色，再复染。

(5) 荧光显微镜观察和图像分析　应用 2～3 段波长光栅的荧光显微镜检测荧光杂交信号。根据染色体的互补性，在每对同源染色体上应当可观察到相同的杂交信号。但被测 DNA 有 DNA 序列的获取时，所获取 DNA 序列的染色体区域对应于对照的 DNA，被测 DNA 的荧光信号增加。如出现 DNA 序列的丢失，则可发现相应的对照组 DNA 的荧光信号增加。应用低温电荷耦合器件相机和合适的光栅，拍下荧光杂交类型的图像并应用相应的电脑软件进行分析，即可得出沿染色体全长的被测和对照的荧光杂交信号的比率。CGH 的技术流程如图 11-4。

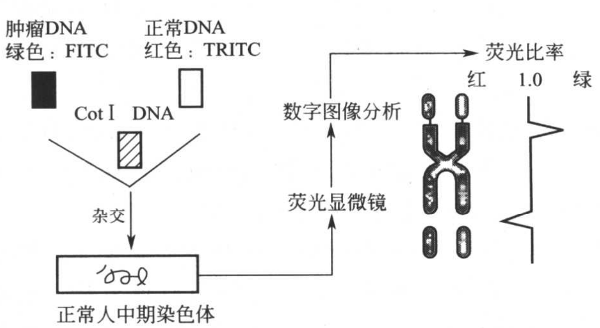


图 11-4　CGH 的技术流程

2. 比较基因组杂交技术的应用

近年，CGH 已经被广泛应用于染色体的获得或丢失的检测。由于它能有效地发现多种肿瘤细胞中染色体的获得或丢失的情况，而这种染色体的变化又与肿瘤的发生和肿瘤细胞的某些生物学行为有关，即与癌基因和抑癌基因在肿瘤的发生、发展过程中的作用有关，所以可以利用 CGH 检测不同肿瘤中染色体扩增、获得和丢失，为应用分子生物学方法在该染色体部位分离出新的基因序列提供新思路。

3. 比较基因组杂交技术的主要优缺点

CGH 与目前常用的一些遗传学分析方法相比，有其明显的优点。首先，该技术既可应用于新鲜或冻存的手术标本，又可应用于石蜡存档标本，还可应用培养的细胞株，而且仅需要很少量的细胞或组织成分。其次，实验方法既简单又可以获得全基因组改变的详尽资料。

但 CGH 也存在一定的缺点和局限性。与其他一些以 PCR 为基础的检测方法相比，它在检测基因的缺失方面敏感性较低，其检测的敏感性还受到待测标本所含正常间质细胞和淋巴细胞数量的影响。CGH 不能发现常见于血液系统或间质性肿瘤中的平衡的染色体异位或重排。CGH 的结果还受到图像分析系统可靠性和转换性能的影响，而且目前这一套设备和配套的分析软件也价钱不菲。

CGH 是一种研究细胞系或生物标本的整个基因组 DNA 获得与缺失的非常有应用前途

的分子细胞遗传学方法，一次实验即可检测出待测标本整个基因组 DNA 拷贝数的增减，亦可在物种间进行染色体同源性比较。

三、微阵列-比较基因组杂交技术

随着生命科学与自然科学的发展以及学科交叉的不断渗入，各种技术之间的联合应用成为趋势。目前，一种将基因芯片和 CGH 相结合的新技术——微阵列-比较基因组杂交技术（microarray-CGH，又称 array-CGH 或 matrix-CGH）日趋成熟，并以其独特的优势而备受瞩目。

1. array-CGH 的特点

array-CGH 实际上就是用微阵列取代传统 CGH 的中期分裂相，使荧光标记的测试 DNA 探针和参照 DNA 探针竞争性地与微阵列上的短片段靶序列杂交。根据被检组织基因组的大小和实验要求，微阵列上的核苷酸靶序列可来源于不同的基因组文库，如 YAC（0.2～2Mb）、BAC（300kb 左右）、P1（约 70～100kb）、PAC（约 130～150kb）和 cosmid（约 30～45kb）等文库。与传统的 CGH 相比，array-CGH 技术在以下两方面具有明显的优势：①灵敏度和精确性，由于染色体上的 DNA 是以高度密集和超螺旋的形式存在着，因此传统 CGH 只有在 DNA 序列缺失达 10～20Mb（细胞株）或 20～30Mb（原发肿瘤）以上或序列扩增时扩增子与扩增拷贝数之积至少达到 2Mb 时才能被检测出来，故传统 CGH 所提供的信息中必然包含为数众多的基因，需要作进一步的精确定位。array-CGH 避开了复杂的染色体结构，所杂交的靶序列仅为包含了少数基因的一段段短 DNA 片段，所以能找出传统 CGH 检测不出的 DNA 序列拷贝数的差异，并同时将扩增或缺失的范围精确地定位在某个或某几个已知基因或 EST 上。②自动化、程序化，染色体带型的复杂性和个体差异决定了核型分析不可能全部实现机械化，必须依靠经验丰富的细胞遗传学家对核型分析软件得出的结果进行校正后才能进一步分析基因组的不平衡性。因此传统的“CGH”技术受人为因素的限制，需要一定的经验技术和劳动力支持。array-CGH 技术中不需要染色体核型的制备和分析，与普通的基因芯片检测表达谱的过程一样，其结果完全可以由机器和计算机自动操纵控制，既快速又直观。

2. 原理与方法

array-CGH 的原理（见图 11-5）与中期染色体 CGH 相同，也是通过用不同颜色荧光素

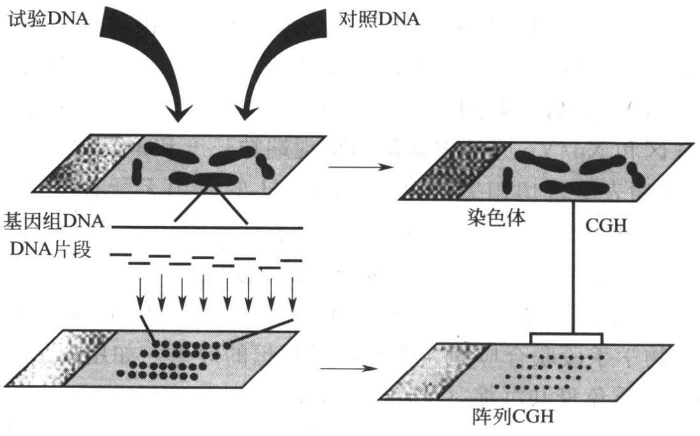


图 11-5　array-CGH 原理

分别标记待测及正常对照 DNA，并同时与制备好的芯片杂交，通过观察红绿荧光的强度比率来了解肿瘤 DNA 拷贝数改变。

array-CGH 与中期染色体 CGH 操作方法存在许多相同之处，甚至可以说基本相同，区别只是在于杂交靶点的不同，array-CGH 操作步骤简要表示如下：芯片制备（根据不同需要）→待测、正常组织 DNA 提取→荧光标记→杂交→洗片→干燥→靶点→再用 DAPI 溶液处理→图像采集、读片、分析。

3. array-CGH 的分类

根据微阵列上所固化的核苷酸的性质，array-CGH 可分为 cDNA array-CGH 和 DNA array-CGH 两种类型。前者即以常规的 cDNA 微阵列为平台，基片表面固定的探针来自待测种属的 cDNA 文库。cDNA array-CGH 技术使得大规模、高通量研究基因剂量的改变成为现实。若与基因芯片技术联合应用则可分别在基因组的复制和转录两个水平上检测基因表达谱和基因不平衡性。DNA array-CGH 与 cDNA array-CGH 不同之处在于 DNA array-CGH 技术是直接将基因组 DNA 片段固化于芯片表面，因此靶序列上不但有基因编码区，还包含了内含子、重复序列、其他调控序列等非翻译区，增加了实验结果的信息量。同时由于基因组 DNA 片段长于 cDNA 片段，因而 DNA array-CGH 的杂交信号强于 cDNA array-CGH，提高了实验的敏感度。但 cDNA array-CGH 芯片制备相对容易，cDNA 片段多介于 0.5～2kb 之间，可很方便地设计引物，通过 PCR 从 cDNA 文库中得到所需的全部序列，而 cDNA 上广泛分布的表达序列标签也有助于对所得序列进一步纯化和点阵。因此这两种 array-CGH 技术各有所长，在实际工作中可根据具体情况和实验要求作出选择。

4. array-CGH 在基因组学研究中的应用

利用 array-CGH 进行肿瘤基因组的研究已成为肿瘤研究中的热点。所采用的芯片既可以是根据研究目的针对某个特定染色体或基因组的一段区域专门设计，也可以根据所研究肿瘤的特点和现有的生物学信息，选取较易产生基因不平衡的区段，或用全基因组芯片对所有已知和未知基因进行筛查。迄今为止，不论采用哪种芯片，实验结果均证实 array-CGH 比传统的 CGH 敏感性更高，可以检测出传统 CGH 检测不到的基因拷贝数异常。基因组的异常可表现为染色体区段异常和基因拷贝数目异常。应用全基因组 array-CGH 研究发现：不同肿瘤组织在基因组异常方面存在很大的异质性。某种类型的肿瘤很可能是由于特定的基因表达的改变或特定的遗传物质不稳定引起的。

与传统 CGH 相比，array-CGH 更利于发现新的癌基因。由于 array-CGH 对基因的定位更加精准，一旦确定某一序列上可能存在候选癌基因，通过基因组数据库就能很容易找到这个基因。人群中 DNA 拷贝数的多态性增加了基因组分析的复杂性。据报道，染色体末端的端粒区，特别是 2q 区和 Xp/Yp 区的末端最易出现端粒长度和重复频率的多态性。研究发现其中某些多态性与疾病表型密切相关，而 array-CGH 的应用无疑有助于剖析复杂的人类基因组多态性现象，从而揭示出这些与疾病相关的多态性基因。

DNA 拷贝数的改变是可能引起基因表达异常的机制之一，而后者是导致个体发育异常和肿瘤的根源。除了全基因组范围内的 array-CGH 外，越来越多的国内外学者开始将 array-CGH 与其他分子生物学手段联合应用，并取得了理想的效果。如用形形色色的 array，包括 CpG 岛 array，与染色质免疫共沉淀（chromatin immunoprecipitation）技术联合应用，为阐明全基因组范围内蛋白质与基因之间的相互作用或与疾病状态之间的关系提供了线索。

总之，随着基因组研究的深入和对各种生物学问题的探索，array-CGH 技术必将得到

进一步发展并被推广到更多的研究领域中去。

四、染色体原位杂交

染色体技术发现距今已有150多年的历史。染色体检测被广泛用于动、植物及人类的细胞遗传学研究。随着染色体分析技术和分子生物学技术的发展，染色体研究范围也不断扩大。染色体原位杂交技术（*in situ* hybridization of chromosome）主要有：①荧光原位杂交技术（fluorescent *in situ* hybridiaation，FISH）；②染色体上寡核苷酸引物原位DNA合成技术（oligonucleotide primed *in situ* DNA synthesis，PRINS）。

FISH具有探针稳定、操作安全，可快速、多色显示多个不同探针的杂交信号等优点。FISH的灵敏性与探针标记方法和检测仪器性能有关，探针标记时掺入的修饰核苷酸比例直接影响杂交信号强度。FISH探针一般采用随机引物法或切口平移法，如将PCR技术引入FISH探针标记，可使其灵敏度提高到0.25kb。应用慢扫描CCD配合影像处理理软件，增强信噪比，有利于检测微弱信号，如应用TSA系统（tyramide signal amplification）能将杂交信号再放大1000倍，可用于单拷贝基因的定位。FISH分辨率大约为1～3Mb，如果应用强变性剂处理染色体，让DNA分子从蛋白质中分离出来，使双链DNA完全伸展并黏附在玻片上，经处理后，分辨率可达1～2kb。还可采用对分裂中期染色体进行显微解剖（microdissect）法以提高分辨率。FISH的另一个特点是可以联合应用地高辛、生物素等多种标记系统，在一次杂交中可检测多种探针在染色体上的位置及探针间的相互关系，即多色FISH或多靶FISH。FISH技术已被广泛应用于肿瘤研究中的基因扩增、易位重排及缺失等的检测，在肿瘤诊断和鉴别诊断、预后和治疗监控等方面都有重要意义。

PRINS可用于染色体特异α卫星DNA标记。PRINS技术的优点是不需克隆基因作探针，且由于杂交反应在前，标记在后，故非特异性背景低；缺点是信号弱、灵敏度低，为克服这一缺陷，Terkelsen等建立了重复引物原位标记技术（repeated PRINS），重复进行PRINS反应，使信号强度明显提高。基于FISH的原理，发展了多色原位标记（multicolor PRINS），可测定2个以上靶序列在DNA分子上的相互的位置，且特异性比FISH还高。肿瘤细胞的染色体变化是一个非常普遍的现象，可分为原发和继发两类。在肿瘤形成的生物学基础方面，原发性的染色体变化与引起肿瘤的直接原因有关，肿瘤细胞中可以发现各种形式的染色体畸变，如缺失、重复、易位、重排、单体断裂及核内复制等；继发性变化主要是肿瘤细胞核型的改变。染色体的检测对于肿瘤的诊断、鉴别诊断、生物学行为判别等方面都重要意义。

五、基因表达分析技术

基因表达分析是基因功能研究最重要的组成部分。基因表达分为转录及翻译两个阶段，转录是以DNA（基因）为模板生成mRNA的过程，翻译是以mRNA为模板生成蛋白质的过程，检测外源基因的表达就是检测特异mRNA及特异蛋白质的生成。所以，基因表达检测分为两个水平，即转录水平上对特异mRNA的检测和翻译水平上对特异蛋白质的检测，以比较不同组织和不同发育阶段、正常状态和疾病状态，以及体外培养的细胞中基因表达模式的差异。传统的技术包括RT-PCR、RNase保护实验、RNA印迹杂交等，目前已发展了多种基因表达分析的新技术，最为常用的包括差异显示PCR（differential display-PCR，DD-PCR）、代表性差异分析（representational differential analysis，RDA）、抑制消减杂交

(suppression subtractive hybridization，SSH)、基因微点阵（gene microarray）以及基因表达系列分析（SAGE）以及实时荧光定量 PCR（real-time fluorescent quantitative PCR，FQ-PCR）等。

（一）基因表达系列分析技术（SAGE）

SAGE 是近几年来发展起来的一种快速分析基因表达信息的技术。它通过快速和详细分析成千上万个 EST 来寻找出表达丰度不同的 SAGE 标签序列，从而接近完整地获得基因组的表达信息。它不但能快速详细地同时分析成千上万个基因转录子的丰富信息，还能发现新基因，因此是基因表达定性和定量研究的一种有效工具，为了解正常作用机制、肿瘤发病机制、信号分子调控机制等提供了基础工具。这使之成为从总体上全面研究基因表达、构建基因表达图谱的首选策略。相比较而言，基因表达系列分析技术综合了其他几个技术的优点，其不仅可以分析组织/细胞的基因表达谱，而且可以发现新的基因，是目前少数几个基于基因表达谱化分析的开放差异基因表达技术之一。

1. 基本特点

SAGE 能够直接读出任何一种类型细胞或组织的基因表达信息，它测量的不是基因表达水平，而是量化能代表一个基因的转录产物的标签。在 SAGE 中标签是一段固定长度的核苷酸序列，3′端比邻一种特定的限制性酶的最 3′端酶切位点，标签的长度一般是 9 个碱基(10～11 个也可)，常用的标签酶是 *Nla* Ⅲ。它以获得的基因组数据库为依据，结合计算机处理系统，识别含有基因表达信息的短片段标签所代表的基因，从而定量分析有机体某一状态下各基因的表达情况，比较不同状态基因表达水平，了解基因在有机体中调节作用和彼此间相互影响。它有三个特点：一是短标签（10～14bp），从每个转录子唯一的位置中获得的这个短核苷酸标签包含有足够的信息，能够唯一确认一种转录物；二是将这些标签连接起来形成一个个长的序列克隆中测序；三是将得到的短序列核苷酸以连续的数据形式输入计算机中进行处理、量化，从而就能对基因转录水平进行分析。它的优点包括：①高度敏感，能大规模进行检测；②检测到所有基因、包括未知基因的定量数据；③避免了扩增时的偏倚；④大量的数据可用于多重比较。

2. 基因表达系列分析方法

SAGE 技术的基本实验流程（见图 11-6）。

① 从特定的组织中分离 mRNA，并以分离的总 mRNA 为模板，加入 5′端连接着生物素的寡聚 dT 为引物，与模板 mRNA 的 polyA 尾巴互补配对，逆转录成 cDNA。

② 用 4 碱基识别位点的限制性核酸内切酶（称为锚定酶，anchoring enzyme，AE）酶切 cDNA，因为大多数 mRNA 都长于 $4^4=256$ 个碱基，所以锚定酶可保证在每一种 cDNA 上都有 1 个切点。

③ 用链亲和素化蛋白小球与 5′端连有生物素的锚定酶酶切片段结合，收集每一 cDNA 上的寡聚 dT 与最近的酶切位点之间的片段。

④ 将收集到的全部 cDNA 末端序列均分为 A 和 B 两部分，分别与含有特定型限制性核酸内切酶识别位点的 A、B 接头通过 AE 黏性末端连接。Ⅲ型限制酶可在距其识别位点下游约 10～20bp 处切割双链 DNA，并形成平头末端。这里Ⅲ型限制酶称作标记酶（tagging enzyme，TE)。接头顺序为：引物 A（或 B）序列＋标记酶识别位点序列＋锚定酶识别位点序列。

⑤ 用标记酶酶切连有接头 A 或 B 的短 cDNA 序列，可在距标记酶识别位点下游约 10～

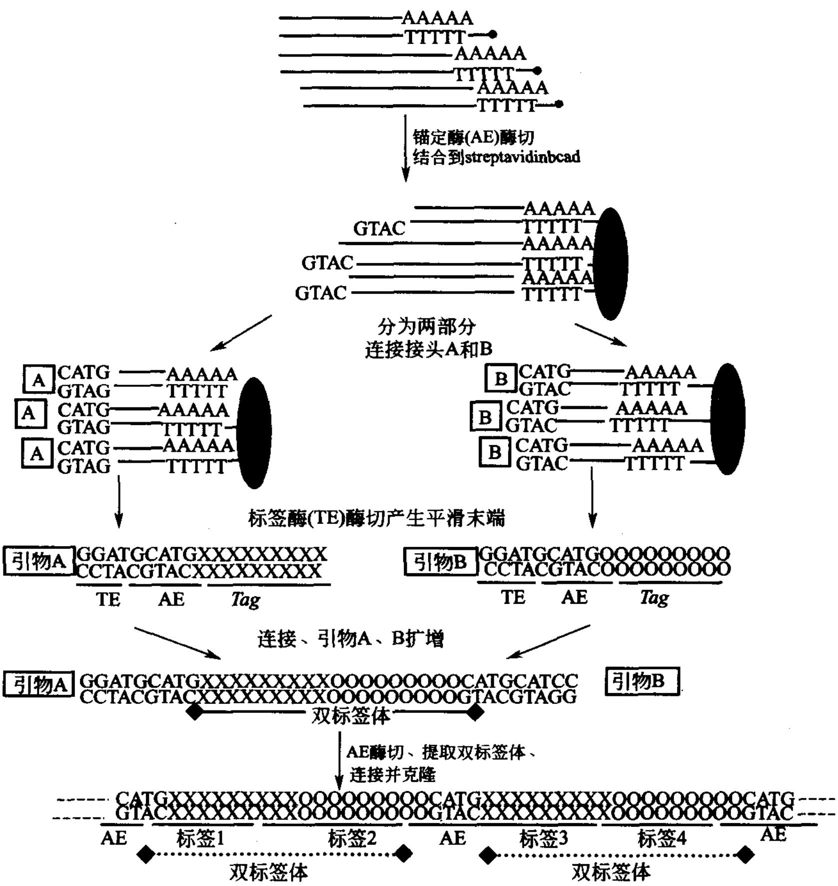


图 11-6　SAGE 技术的基本实验流程

15bp 处切断短 cDNA。将 A、B 两部分 cDNA 酶切片段混合并与末端连接，即构成了左右两端分别为 A、B 接头、中间由来自两个不同 cDNA 末端各 9bp 长的双标记探针。以 A 和 B 为左右引物对双标记探针进行 PCR 扩增，即可得到大量的双标记探针。

⑥ 用锚定酶切割扩增的双标记探针，可切去 A、B 接头序列，将酶切后的双标记探针通过 AE 黏性末端连接，可构成由大量不同的双标记探针连接而成的 DNA 片段，克隆此片段并测序。

⑦ 通过软件对标签数据进行处理。统计原始序列中每个标签序列、数目及发生率，生成相应的报告和丰度指标。

⑧ 分析整理数据，比较 SAGE 文库标签的丰度，分析其意义。

3. SAGE 在医学研究中的应用

（1）应用于人类和动物基因组研究　1995 年 Velculescu 等选择 *Bsm* FⅠ和 *Nla*Ⅲ分别作为标签酶和锚定酶，对 1000 多个标签进行手工测序从而发现了胰腺功能的基因表达特性，在分析的 1000 个标签中，95%以上的标签能够代表唯一的转录物。

（2）SAGE 也能应用于寻找新基因　如果一个标签检索已知序列时没有同源序列，13 碱基片段（标签 9 个碱基加上锚定酶的位点序列 4 个碱基）就可作为探针筛选 cDNA 文库得到 cDNA 克隆。肿瘤基因组计划（cancer genome anatomy project，CGAP）是一个致力于

解码肿瘤形成和进展中基因变化的肿瘤研究联合网络组织，它综合了包括 SAGE 在内的多种高效分析手段，提供人类和小鼠转录子的全面信息，提供了 SAGEmap 这样一个定量的数据库，包括哺乳类动物基因采集（mammalian gene collection initiative）、人类基因系列的单核苷酸多态性（SNPs）研究、肿瘤细胞染色体畸变（cancer chromosome aberration project）及染色体重排（chromosomal rearrangements）等项目。

（3）用于定量比较不同状态下的组织细胞的特异基因表达　Reprogen 等通过分析 9 个人类卵母细胞的 50000 个 SAGE 标签，发现与已知基因同源的序列主要是表面受体、第二信使系统、细胞骨架、凋亡或分泌相关蛋白等，发现了许多以前未知在卵母细胞表达的蛋白的编码基因，这些有助于确定控制体细胞核全能性分化的因素。

（4）衰老机制研究　Wells 等利用 SAGE 及寡核苷酸微点阵分析技术分析了 18～24 岁和 66～77 岁各 8 人的肌细胞 RNA，重点对发现的在 5 个月和 30 个月小鼠的肌细胞中表达水平不同的 61 个基因（Lee CK 等）进行了检测，发现 17 个基因存在与鼠类相似的年龄相关性，而对于 32 个基因，年龄对其表达水平的影响与鼠类不同。在年老的鼠类中有 10 个应激反应基因表达增高，而在老年人肌细胞中，发现其中有 5 个基因表达未增高。

（5）St Croixt 等为了弄清肿瘤血管生成的分子机制，比较了分别从同一个病人的正常结肠黏膜和结肠癌组织中分离纯化的血管内皮的 96000 个标签，这些标签代表了 32500 个唯一的转录物，在优势表达的 170 多个转录物中，79 个存在差异表达，其中 46 个基因在肿瘤相关上皮中特异地性表达，包括至少 7 个基因编码细胞外基质形成和重塑蛋白。这些肿瘤上皮标志物大多表达在众多肿瘤类型中，同时也在正常组织创伤愈合、黄体形成（corpus luteum formation）时存在。这些研究证明了 4 点：①正常组织和肿瘤上皮高度相关，共同表达许多上皮细胞特异性的标志物；②肿瘤来源的上皮与同类型的正常组织上皮不同，也与原发的肿瘤上皮培养物不同（46 个高表达的基因在体外培养中检测不到或表达极低）；③这些基因特征性表达在几种不同组织类型来源的肿瘤，肿瘤上皮和正常上皮细胞在分子水平显著不同；④肿瘤上皮标志物（tumor endothelial markers TEM1 到 TEM9）同时也在正常组织创伤愈合、黄体形成时存在表达。

（二）RNase 保护试验（RPA）

同位素或生物素标记的 RNA 探针可以像 DNA 探针一样用于分子杂交实验，从而对特异性 RNA 含量进行测定并可对基因和 mRNA 的结构进行分析。RNase 保护试验（RNase protection assay，RPA）就是一种利用 RNA 探针对基因表达产物 RNA 进行定性定量分析的有效方法。利用 RNaseA 和 RNaseT1 能专一性降解单链 RNA 而双链 RNA 受到保护的特性，用体外转录合成的放射性标记的 RNA 探针与待检测 mRNA 进行液相杂交，使 RNA 探针与待测 RNA 间在互补序列处形成 RNA* ：RNA 杂交体（* 表示放射性标记），然后用 RNase A 和 RNase T1 切割此 RNA 杂交体，受到保护的 RNA 通过聚丙烯酰胺凝胶电泳分析，以探针上的标记物采用相应的方法可以分析出受到保护的 RNA 片段的长度和含量，从而对待测 mRNA 进行定性和定量分析。此法可用于 mRNA 定量、mRNA 末端定位、mRNA 剪切部位及确定内含子在相应基因中的位置。

六、模式生物体研究

随着人类基因组测序工作及精细图绘制的完成，大量新基因序列随之产生，如何理解这些基因的功能是摆在各国科学家面前的一项艰巨任务。鉴定基因功能最有效的方法就是观察

基因表达被阻断或增加后在细胞和整体水平所产生的表型变异，而模式生物体（model organism）提供了一个强有力的解析基因功能的工具，可以通过寻找与人类某个未知功能基因相当的模式生物体内的基因（即所谓直向同源物），并对其功能进行研究，从而加深对人类该基因功能的认识。

通过对模式生物体表型变化（突变体）的分析，可在动物整体、细胞或分子水平认识基因的功能。模式生物体是全面准确解释基因功能的有效手段，将被更为广泛地应用于未知基因的功能和人类疾病的研究。迄今为止，已有许多生物被作为模型，如大肠杆菌、酿酒酵母、秀丽线虫、果蝇、河豚鱼、斑马鱼、小鼠、曲霉菌（*Aspergillus nidulans*）等。但线虫和果蝇与人相距甚远，包括造血功能在内的许多人体重要生理功能以及相关的疾病基因，在这些模式生物体内并不存在。由于小鼠的基因组序列已完成，且与人类有很高的同源性，是目前很好的一个模式生物体。对小鼠不同状态和突变体基因的表达的研究对人类同源基因的功能研究来说是一个很好的资源，但其繁殖速度较慢，且体积相对较大，表型出现的时间也相对较长，因此限制了其作为大规模高通量新基因功能研究的模式生物体的应用，而斑马鱼可以弥补这些不足。

国外 20 世纪 80 年代以来，继成功建立转基因动物技术体系后，结合基因同源重组和胚胎干细胞培养技术建立和发展了基因敲除（gene knock-out）技术，使基因的定位敲除、精细突变的引入、染色体组大片段的重排和删除、组织特异性或时相特异性基因敲除均已成为现实。这些技术的建立和应用极大地促进了生命科学研究的发展，为未知基因功能和人类疾病发病机制的研究提供了强有力的技术手段。

（一）转基因动物

1. 概念

转基因动物是指染色体基因组中整合有外源基因并能遗传给后代的一类动物。整合到动物染色体基因组的外源基因被称为转基因。转基因技术则是指制备转基因动物所需的一套技术，它涉及外源基因的组建、载体、受体、基因导入技术、供转基因胚胎发育的体外培养系统和宿主动物等许多方面。

当制备转基因动物时，外源基因可能只整合到动物的部分组织细胞的基因组，也可能整合进动物所有组织细胞的基因组中。只有部分组织细胞的基因组中整合有外源基因的动物，被称为嵌合体动物（chimeric animal）。只有当外源基因整合人的生殖细胞时，才能将其携带的外源基因遗传给子代，否则外源基因将不能传给子代。一般用胚胎干细胞法、逆转录病毒载体法制备的第一代转基因动物均为嵌合体动物，而显微注射法得到的第一代转基因动物中，也约有 20%为此类动物。如果动物所有的细胞均整合有外源基因，则具有将外源基因遗传给子代的能力，通常把这类动物称为转基因动物。

2. 基本原理

转基因动物的基本原理是将改建后的目的基因（或基因组片段）用显微注射等方法注入实验动物的受精卵（或着床前胚胎细胞），然后将此受精卵（或着床前胚胎细胞）再植入受体动物的输卵管（或子宫）中，使其发育成携带有外源基因的转基因动物，人们可以通过分析转基因和动物表型的关系，从而揭示外源基因的功能；也可以通过转入外源基因培育品种优良的工程动物等。动物的转基因技术按照外源基因导入的方法和对象的不同，主要有显微注射技术、胚胎干细胞（ES 细胞）技术、逆转录病毒载体技术三种方法。

（1）显微注射法　1980 年和 1981 年 Gorden 等报道了用显微注射法将外源基因注入小

鼠受精卵中得到了转基因小鼠。本法的特点是通过显微注射直接将外源基因导入受精卵，再移植到受体，使其发育成转基因动物的一种方法。这是应用最早和成功率较高的一种方法。目前，本法也是最常用的一种方法。优点：①外源基因在宿主染色体上的整合率相对较高（和其他方法比较）；②以此种方式导外源基因时，无需载体；③在60%～80%的转基因子代鼠中，外源基因分布在所有的组织和细胞中；④导入的外源基因可长达50kb。不足之处：①外源基因的整合是随机的，整合率不能控制；②有些动物的胞核看不清楚，如猪、山羊等，须经特殊处理，才能有效导入；③需要较精密的仪器，费用昂贵。

（2）逆转录病毒法　1974年Jaenish和Minti首先将SV40病毒DNA注入到小鼠胚胎的囊腔中，此后就开始（1981年）了用逆转录病毒作载体将外源基因导入到胚胎的研究。

逆转录病毒具有侵入宿主细胞并整合于细胞染色体DNA的能力。将插入有外源基因的逆转录病毒载体DNA，通过辅助细胞包装成高感染滴度的病毒颗粒，再感染桑椹期的胚胎细胞，随后将胚胎导入子宫，可发育成携带外源基因的子代动物。原则上载体中可插入各种基因组DNA和cDNA。逆转录病毒载体的容纳量在10kb以下，不能插入较大的外源基因。优点：①感染率高，且胚胎存活率高；②病毒DNA随机单拷贝整合；③宿主范围广。不足之处：①整合位点是随机的，其整合率不高；由于病毒感染过程是在多次卵裂之后，外源基因很难整合于所有胚胎细胞中，故多数子代动物为嵌合体；②重组的逆转录病毒不够稳定，外源基因易发生重排或丢失；③病毒载体的容量不大；④病毒DNA序列可能会干扰外源基因的表达。

（3）胚胎干细胞法　20世纪80年代中期人们开始研究用胚胎干细胞建立转基因动物的方法。该方法将外源基因直接导入胚胎干细胞，经体外培养筛选后，通过显微操作将其注射到囊腔中，再移植入受体的子宫内，使其发育为嵌合体的转基因动物。

胚胎多能干细胞（embryonic stem cell，ESC，ES；或evanskaufmancell，EK）是指从囊胚期的内细胞团中分离出来的尚未分化的胚胎细胞，能在体外培养，具有发育全能性。它具有胚胎细胞和体细胞的某些特性，既可进行体外培养、扩增、转化和制作遗传突变型等遗传操作，又保留了分化成包括生殖细胞在内的各种组织的潜能。可以利用ES细胞转基因技术和嵌合体技术得到转基因动物。优点：①可用多种方法将外源基因导入ES细胞，其细胞的鉴定和筛选比较方便；②可预先在细胞水平上鉴定外源基因的拷贝数、定位、表达的水平及插入的稳定性等，故转基因动物外源基因整合率的可控制程度高于微注射法；③ES细胞注入囊胚以及囊胚移植到子宫，在操作上比较简便。不足之处：①本法中建立ES细胞系的本身就是一种高难技术，其培养条件至今无突破性进展；②ES细胞的保存和维持也不容易；③所得个体为嵌合体，如发生性嵌合，将导致生殖机能紊乱和不育；④所需时间较长。

除上述3种方法外，精子载体法已引起世界各国科学家的极大兴趣，但在实践中还存在分歧，这一技术尚处在完善和发展阶段，这里不再介绍。

3. 嵌合体动物

嵌合体动物（chimera）指的是两种（或以上）来源不同细胞群体构成的个体。至于来源于一个受精卵的体细胞突变造成遗传上不同的两类细胞群体而构成的个体叫做镶嵌体（mosaic），它与嵌合体是有区别的。

嵌合体动物是由遗传上不同的两类胚胎或胚细胞组合形成的。因为是由两种细胞群体构成的器官和组织，其好处是能在同一个个体条件下观察细胞间的相互作用。哺乳动物在其发育过程中分裂的细胞是在运动中分化，所以细胞间相互作用在形态形成上起着重要作用。利

用适当的基因标志可以分析出细胞间相互作用和细胞的谱系。

4. 转基因动物模型在医学研究中的应用

（1）在免疫学研究中的应用　转基因动物技术对于探索免疫潜能细胞（immunocompetent cell）的发生和体内调节是一非常有用的工具，免疫球蛋白基因在转基因鼠中可被表达，充分显示了转基因鼠技术在富含特定抗体的B细胞种群研究中的作用；此外，转基因动物技术在免疫球蛋白的多样性，即免疫球蛋白基因的重排方面具有独特的优点，是其他任何实验方法所不及的。1983年，Brinster等将一功能性重排的小鼠免疫球蛋白（Ig）κ基因注入小鼠受精卵，建立了转基因鼠。他们的研究结果显示，重排的Ig基因包含了所有的组织特异性表达所需信息，即不管导入的功能性重排Ig基因整合于小鼠染色体的什么位点，在连续的传代中，都能获得稳定和正常的表达；他们的结果还揭示，导入的功能性κ轻链基因控制内源性轻链基因的重排能力与转基因的高水平表达有关，这说明产生轻重链的数量对于有效的等位排斥是有影响的。通过把鸡和兔κ基因片段导入小鼠，可研究免疫球蛋白基因片段重排的组织特异性，结果表明这些轻链转基因不仅能在B细胞中重排，而且在胸腺中也能重排，而正常情况下胸腺是不发生κ基因重排的。因此，这种小鼠表达了杂交的Ig分子，这提示在遗传工程动物中产生种间单抗是可能的。

（2）在病毒性疾病研究中的应用　主要有以下几个方面：①将病毒的全基因组DNA导入小鼠基因组。这种转基因小鼠不仅具有表达病毒蛋白能力，还能进行病毒DNA的复制。目前已制备成功的有AIDS病毒、乙型肝炎病毒和人白细胞白血病Ⅰ型病毒的小鼠模型。它们在发病机理探讨和治疗药物的筛选等方面具有重要价值。②将病毒特定基因如乙肝病毒X基因、S基因等导入小鼠基因组，以了解特定基因表达产物的致病机理。③将特定病毒感染宿主的受体基因导入小鼠基因组，使转基因小鼠对特定病毒易感，可用于治疗药物筛选和疫苗效果检测等，如将人脊髓灰质炎病毒受体基因导入小鼠，获得对该病毒易感的转基因小鼠，可替代猴用于检测口服脊髓灰质炎病毒疫苗的毒力以及病毒致癌的分子机制研究。④抗病毒感染的宿主基因导入转基因小鼠，以提高抗病毒能力，如具有常染色体显性（Mxl^{+}）等位基因的小鼠能抵抗流感病毒的感染，Mxl^{-}小鼠均死于流感病毒的感染。将受自身启动子控制的Mxl基因导入小鼠，在注射干扰素、双链RNA或病毒后，转基因小鼠Mxl启动子被激活，Mxl蛋白表达，结果转基因小鼠能抵抗高达5000个致死剂量的流感病毒攻击。

转基因小鼠在肝细胞损伤免疫机理研究中也有独特的用途，F. V. Chisari发现产生的所有转基因小鼠在临床和组织上均表现正常。HBsAg阳性的转基因小鼠用HBsAg加上弗氏完全佐剂不能诱导产生特异性抗体，而HBsAg阴性的转基因小鼠则产生免疫反应。Babinet发现虽然转基因小鼠产生HBsAg，但在6个月内没有任何病理变化，表现为一种持续的带病毒状态。这一结果验证了这样一个假设，乙型肝炎病人的肝细胞损伤不是由HBV HBsAg表达直接引起，而是通过对肝细胞膜上的病毒抗原发生免疫反应造成。因此，可以用转基因小鼠模型来研究免疫耐受与肝细胞损伤的关系以探讨发病机理。此外，转基因小鼠还为研究第Ⅰ和第Ⅱ类主要组织相容性抗原的功能提供了新的手段。

（3）在遗传病研究中的应用　采用转基因技术，可以人为定向地突变小鼠基因组中的某个特定基因，也可以人为地加入一个或多个外源基因，因而在制备遗传病的动物模型方面有很广泛的应用前景。例如，将功能缺陷的基因导入小鼠生殖细胞，制备得到遗传性疾病的转基因小鼠模型。透明软骨胶原纤丝有Ⅱ型、Ⅵ型、Ⅺ型三个型；Ⅱ型胶原蛋白由al（Ⅱ）基因编码的肽链同源三聚体构成，Ⅱ型胶原突变得到显性表型，已制备出五种Ⅱ型胶原突变型

基因的转基因小鼠，表现为肢体短小的侏儒表型，是显性遗传病模型的典型例子。通过 EK 细胞导入基因发生重组，筛选插入突变型，制成了隐性遗传病模型。

应用反义 RNA 技术原理，将针对特定基因的反义 RNA 的转录质粒作为目的基因导入小鼠，制备该基因缺陷型的遗传病转基因动物的模型。这一新颖的思路已于 1988 年由 Katsuki 等的工作获得实现，他们成功地制成了髓鞘发育不全症（MBP 缺失型）的转基因小鼠模型。目前，已制成的模型有 30 多种，最近制成的有镰形细胞贫血、囊性纤维化和肾功能衰竭模型等。据估计，在几年内，遗传病的动物模型数量还将急剧上升。这些模型的建立在观察转基因动物体内的表达特性及表达产物的生理学效应、遗传病的发病机理以及基因治疗试验等方面有研究价值。

（4）基因治疗方面　基因治疗（genetherapy）是通过改变患者遗传物质从而达到治疗疾病的一种方法。改变患者细胞遗传物质的主要途径有：①引入功能正常的基因；②置换致病基因；③修饰致病基因。第一种方法相对简单，只需将目的基因转入体外培养的患者细胞，然后再将其移入患者体内（或用特定的方法将目的基因直接注入体内患病组织使其进入细胞），从而达到治疗疾病的目的；后两种方法则难度较大，需要通过同源重组等技术对患者细胞的基因组 DNA 进行定点整合（site-directed integration）或修饰，才能替代或校正致病基因，达到治病的目的。基因治疗的方式包括体细胞（somatic cell）的基因治疗和生殖细胞的基因治疗两种，相比较而言，体细胞的基因治疗简单易行，目前除有大量的动物体细胞基因治疗的资料外，人的基因治疗也已开始，到 1993 年止，已有 50 种以上的疾病用于基因治疗或实验室研究，且正在以较快的速度递增，但存在目的基因的表达不能有效地控制，目的基因的靶向性有待提高等问题。另外，体细胞基因治疗最根本问题是不能根治疾病，而生殖细胞的基因治疗则为根治疾病提供了有效的方法，其基本原理是在生物发育的胚胎期（包括受精卵）即对其遗传物质进行修饰，从而达到校正基因表达和治疗疾病的目的。尽管这种基因治疗技术目前还不能用于人类，但已广泛用于动物疾病的基因治疗，其中包括生殖系统的疾病、心血管疾病、肌营养不良等。这主要表现在两个方面：一方面是转基因技术直接用于生殖细胞的基因治疗，如小鼠β-地中海贫血、小鼠不育症、肌营养不良症；另一方面是转基因动物可为基因治疗提供模型，如心血管疾病转基因动物模型与基因治疗、神经系统疾病动物模型的建立、肺囊性纤维化转基因动物模型的建立与基因治疗。

原则上，生殖细胞基因治疗的最大优点是可以根治遗传性疾病，但实际上由于其治疗的方式是将外源功能性的基因导入动物的生殖细胞中，因此同转基因动物的建立一样，存在着许多问题，如插入突变可能会导致其他疾病的产生，外源基因是否表达、表达如何调控以及表达在传代中能否维持等。由于生殖细胞的基因治疗难度大、操作复杂以及存在着一些社会问题（如伦理、道德等），所以目前生殖细胞的基因治疗只限于动物实验，但一些专家预言，随着胚胎实验技术的提高，人类生殖细胞的基因治疗是完全有可能的，现有的和正在开展的动物水平的研究将为人类的基因治疗提供更多的第一手资料。由于简单遗传病的病因较清楚，所以目前动物生殖细胞的基因治疗尚多属于单基因缺失或突变的遗传病，相信随着对病因的遗传学分析和对疾病候选基因功能以及发病机制的深入研究，多基因疾病的生殖细胞的基因治疗也将成为现实。

（5）在肿瘤学研究中的应用　近 10 年来肿瘤学研究最大的突破是发现了癌基因，到 1990 年，已发现的癌基因约有 100 个，其功能各不相同。研究证明各种脊椎动物都携带有癌基因，在正常情况下引起细胞癌变，只有在某些条件下才能被激活，致使癌基因发生扩

增、癌基因中甲基基团的丢失等变化而导致细胞癌变。癌基因的这些研究首先是借助转基因技术发展起来的，如对 SV40 的大 T 抗原基因和 *ras* 基因的致癌性已有不少研究。因此，建立携带有肿瘤基因的转基因动物是癌基因活性和肿瘤发生研究的一种极为重要的新方法，它具有四维特征，能够从空间和时间角度同时观察目的基因的活动情况。

通过对乳腺癌转基因小鼠模型研究表明，肿瘤生长与正常细胞同样受内源激素组织内的多肽生长因子的调控。乳腺管上皮细胞的异常增生是由于癌基因和肿瘤抑制基因等的遗传改变，才导致更多的恶性细胞的克隆的增长，形成乳腺癌。转基因动物的研究还支持了肿瘤发生至少需要 2 次转化的肿瘤多步发生假说，如通过转 C-mye 和 SV40。基因的实验研究结果表明，转入的 C-mye 和 SV40 基因在动物各个器官都能表达，但仅有少数几个器官会发生癌变，这就有力地支持了肿瘤多步发生的假说。

利用转基因技术，制备某些高频率发生肿瘤的转基因动物，可以加深对肿瘤发生的了解，由这些转基因小鼠，我们可以分析某一确定组织的癌前状态，分析研究肿瘤的发生发展过程。利用这些处于癌前状态的转基因动物，还可为有关肿瘤的多步发生中“第二步打击因子”的实验提供了强有力的手段。将不同的癌基因再导入这些处于癌前状态的动物，可以证实它们对肿瘤发生发展的不同影响，并有可能为抗癌治疗建立更多可靠的动物模型。因此，转基因技术在肿瘤的病因学、发病学和治疗学的应用研究中具有十分重要的意义。

(6) 在心血管疾病研究中的应用 转基因小鼠在理解冬种调节心血管功能因子方面将起重要作用。转脂蛋白运输基因小鼠，可以用来了解这些基因在体内的脂蛋白代谢中的作用。不论是单个还是多个基因的转基因小鼠对人脂蛋白紊乱和动脉粥样硬化的研究都是有用的模型。转纤维蛋白溶酶原激活抑制-1(PAI-1) 基因小鼠会特异性地发生静脉闭塞，另外已经建立了有突发性高血压的转基因小鼠模型。

(7) 动物生物反应器 将有生物医学重要应用价值的生物活性蛋白质基因导入家畜或家禽受精卵，在发育成的转基因动物体液或血液、乳、腹水和尿中收获基因产物，这就是所谓的“动物生物反应器”，在此系统中动物就像一个活的发酵罐，除温度、pH 和氧气由动物本身自动控制外，本质上与微生物发酵系统并无差别，都是通过生物的基本单位——细胞来表达外源基因产物，由此很可能发展成新型的制药工艺——分子农场 (molecular farming)，或基因农场 (gene farming)。

一种理想的动物生物反应器除要求产率高外，还要求分泌产物所在环境中成分简单，便于提纯。产物除了可被诱导和控制外，应不影响动物的正常代谢，且能被分泌到体外。乳汁为最理想的分泌部位。tPA（组织型溶纤蛋白元激活因子）是治疗血栓的理想药物，Westphal 等将该蛋白的基因修饰后导入小鼠受精卵，结果在小鼠分泌的乳汁中检测到了人 tPA。如能在转基因山羊奶中生产这种乳汁分泌的 tPA，则只要培养出 300 头这种羊，其奶即可供应全世界心脏病人使用。如用奶牛表达。则数 10 头即可够用。除 tPA 外，凝血因子Ⅳ和 α-抗胰蛋白酶也在转基因绵羊中得到了表达。β-乳球蛋白是反刍动物乳汁中的主要蛋白成分，而在啮齿动物中却不存在，Simons 等将 β-乳球蛋白导入小鼠受精卵，得到的转基因小鼠表达了丰富的羊奶主要成分：β-乳球蛋白。因此，转基因动物作为“生物反应器”，应用前景广阔，人类已完全有能力按照需要利用它生产乙肝病毒抗原、卵泡刺激素、促黄体生成素、凝血因子、组织溶纤维蛋白原激活因子等生物制剂和生物活性物质，还可改变奶制品的成分，为医药、食品及畜牧业的发展开辟了广阔的天地。

综上所述，转基因动物这种生物高新技术可以用来建立人类疾病的各种转基因动物模

型，用以研究外源基因在整体动物中的表达调控规律，从而对人类疾病的病因、发病机理和治疗学，将会起到极大的促进作用。另外，还可利用这种高技术改变动物基因型，使其表现型更符合人类需要；亦可用转基因动物产生人类所需的生物活性物质。目前，至少有 51 家公司和 141 个机构参与转基因动物的研究与开发。随着研究的不断深入，转基因动物必将渗透到生物医学的各个领域，对生物医学的研究起着巨大的促进作用。

（二）基因打靶技术

基因打靶（gene targeting），亦称为基因定点同源重组（site-specific homologous recombination），是 20 世纪 80 年代发展起来的一项重要的分子生物学技术，是利用基因转移方法，将外源 DNA 序列导入靶细胞后，通过外源 DNA 序列与靶细胞内染色体上同源 DNA 序列间的重组，将外源 DNA 定点整合入靶细胞基因组上某一确定的位点，或对某一预先确定的靶位点进行定点突变，从而改变细胞遗传特性的方法。基因打靶最常用的策略是通过同源重组使特定靶基因失活，以研究该基因的功能，称基因敲除（gene knock-out）；而基因敲入（gene knock-in）则是通过同源重组用一种基因替换另一种基因，以便在体内测定它们是否具有相同的功能，或将正常基因引入基因组中置换突变基因以达到靶向基因治疗的目的。通常所说的同源重组、基因敲除及基因打靶三个名词是有联系的，确切地说，基因打靶是指将外源基因通过打靶载体以同源重组的方式定点整合到基因组中的指定位置，多是 gain of function（获得功能），而基因敲除也是通过打靶载体将某个指定的通常是内源基因表达沉默，即使之 loss of function（丢失功能），一般我们将二者等同。在人类基因组序列分析的结构基因组学的研究完成后，基因打靶将成为在功能基因组学研究中最直接和最有效的方法之一。

1. 基因打靶技术的原理

进行基因打靶，首先要设计和合成一个将要导入靶细胞的靶向载体（targeting vector）。该载体不仅含有需要插入的 DNA 序列，其两端还含有与靶基因座上的序列相同的核苷酸片段，即同源重组指导序列。将此载体用基因转移的方法导入靶细胞，通过外源载体和内源靶位点相同的核苷酸序列之间的同源重组（homologous recombination，HR），使外源 DNA 定点整合到靶细胞的特定的基因座上。关于同源重组的分子机制目前尚未阐明，但现在已相继提出了多种解释同源重组的模型，其中 Meselson-Radding 模型不仅可以解释交互重组现象，而且还可以圆满地解释基因转变现象，所以被广泛接受。

2. 基因打靶的操作程序

（1）打靶载体的构建　把目的基因和同源重组指导序列构建到带有标记基因的载体上。基因同源重组的发生依赖于同源序列的长度，目前认为 30～40bp 的同源序列长度将是发生同源重组的最低限度。实验表明，在哺乳动物细胞内，当同源序列长度在 295bp～1.8kb 之间时，重组率与同源序列长度呈正比，当同源序列长度在 200bp 以下时，重组效率明显降低。这对设计打靶载体时选择合适的同源重组指导序列长度具有重要的指导意义。一般情况下，同源重组指导序列为基因组 DNA，而不用 cDNA，以免造成基因组缺失或其他改变。

根据同源重组时载体插入基因组的方式，可将打靶载体分为插入型载体、替换型载体和治疗型载体三种类型。前两种载体主要用于基因的去除或添加，其区别在于载体 DNA 同源序列双链断裂位点的位置不同。插入型载体的断裂位点在同源重组指导序列内，目的基因可在载体的任何位置，载体与染色体靶位点进行一次交换完成同源重组，其结果是整个载体整合到染色体靶位点上；置换型载体的断裂位点在同源重组指导序列的两侧或外侧，目的基因

在同源重组指导序列内，同源重组指导序列与染色体靶位进行两次交换完成同源重组，重组的结果是只有载体的同源重组指导序列及其以内的部分取代染色体的靶位序列，同源重组指导序列以外的部分被切除。治疗型载体主要用于对因突变而丧失正常功能的基因进行纠正。

根据基因打靶目的的不同，打靶载体可有不同的设计方法。基因打靶载体一般包括5′和3′侧的同源序列、两者之间的*neo*（新霉素磷酸转移酶基因）、同源序列外侧的*tk*（胸苷激酶基因）及载体骨架。基因打靶的示意模式见图11-7，其中A表示基因替代型，B表示基因插入型，C表示基因治疗型；a、b、c、d、e与a_1、b_1、c_1、d_1、e_1分别代表同源序列；x代表目的基因片段；c*代表突变的c基因；黑影表示目的基因与和靶基因两端特异片段同源的基因序列；空白表示打靶载体上的基因序列。此外，有几点说明如下：①在A中，由于bxd两端是同源序列a和e，同源序列交叉互换，bxd取代了$b_1c_1d_1$；②在B中，同源序列ab、cd进行交叉互换，导致目的基因r的加入；③在C中，正常基因c连同同源基因b、d，由于a、e交叉互换作用，c取代了c*。

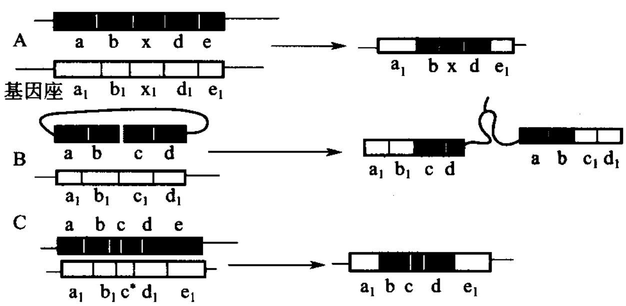


图11-7 基因打靶载体模式图

（2）转化受体细胞 由于在同源序列附近的DNA有双链断裂能促进同源重组，因此在完成打靶载体的构建后，以限制酶线性化载体并除去质粒部分，用电穿孔、显微注射、DNA-磷酸钙共沉淀、脂质体包装和PEG介导等方法把上述重组DNA片段转入受体细胞核内。

（3）筛选 用选择培养基筛选已击中的细胞，通常筛选使用正、负选择法。构建载体时，在靶基因的同源序列中插入正选择标记，在同源序列之外的3′端，甚至3′和5′两个末端接上负选择标记。转化受体细胞，如果所导入的重组DNA与受体细胞基因组DNA之间发生非同源重组，则外源基因通常是从头至尾均整合进入受体细胞基因组中，此时正、负筛选标记基因同时表达，若为同源重组，外源目的基因及正选择标记基因会整合到受体细胞基因组同源序列的基因座上，而位于同源序列外端的负选择标记基因则在重组后丢失，因而此时仅有正筛选标记基因表达。常以新霉素抗性基因（neo^r）为正标记基因，以单纯疱疹病毒的胸腺嘧啶核苷激酶基因（*HSV-tk*）作为负筛选基因。以药物G418和GANC（丙氧鸟苷）作双重选择。其中，neo^r基因的产物使细胞具有G418抗性，而*HSV-tk*基因的产物则使丙氧鸟苷变成毒性核苷酸，从而致死。用G418筛选所有能表达neo^r基因的细胞，然后用GANC淘汰所有*HSV-tk*正常表达的细胞，剩下的细胞为命中的细胞。正、负选择载体的筛选流程图见图11-8。

（4）鉴定 观察被击中细胞的生物学特性，并进行分子生物学检测。可以用特异PCR方法鉴定，PCR引物一端以基因组DNA的特定基因座为模板，另一端以导入的外源基因为模板，这样保证扩增后的片段为同源重组产生的特殊片段。对经PCR鉴定的克隆用South-

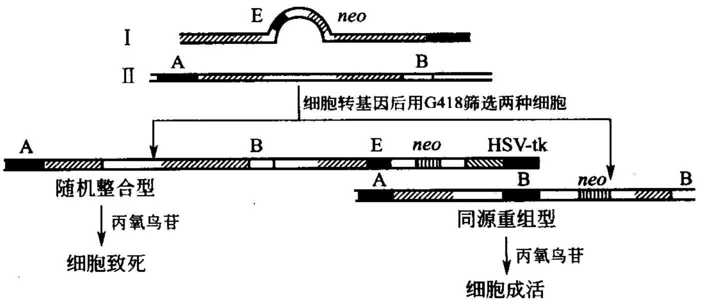


图 11-8　正、负选择载体的筛选流程图

ern 杂交产生特殊带谱的方法来进一步确定同源重组克隆。

3. 提高基因打靶效率的途径

目前基因打靶技术存在的问题是命中率低。虽然从原理上讲，细胞内同源重组要比非同源重组容易发生，但由于基因组上非同源基因座相对于同源基因座要多得多，使得在重组时随机整合的频率要远远地高于同源重组。但基因打靶的概率不是一个固定的量，可以随遗传环境不同而不同。下面介绍一些可以提高基因打靶效率的途径。

（1）基因表达的瞬时改变　通过瞬时改变细胞内基因的表达情况，过量产生某些蛋白，或抑制某些蛋白的产生，来提高基因打靶并降低随机整合。目前对于同源重组的研究提供了几个有希望的候选基因，这些基因的表达水平的改变能提高基因打靶的效率。但控制这些基因表达水平的短期改变很关键，因为持续改变这些基因的表达很可能对基因组有害。

（2）打靶载体的修饰　通过修饰打靶载体可提高基因打靶效率或降低随机整合。目前有两种方法可以使用。第一种方法是在线性化 DNA 的 3′端加入双脱氧核苷酸。这种修饰能以两种方式增加靶位同源重组概率。首先这种方式防止了转染的 DNA 间的连接，由此可减少多拷贝整合到靶基因上，其次它可以加强正负筛选的功能，因为它可以保护末端的负标记基因，防止它降解。第二种方法是在载体的末端造成几百个核苷酸长的单链尾巴。

（3）转化方法　早期使用显微注射法，靶位重组的概率为 1/10～1/20。但这种方法需要专门的设备和技能，因而很快被大量提呈载体的方法所取代。目前，最常用的是电穿孔法。使用大量提呈载体的转化方法，要能向每个细胞中转入非常少的载体分子才能得到最好的打靶概率。在目前的方法中，电穿孔是最接近这一目标的方法。

（4）基因特异性损伤的引入

① 嵌合性核酸酶　由限制酶和稀有位点的核酸内切酶如 I-*SceI* 和 PI-*SceI* 引入的染色体双链断裂可使基因打靶的效率提高 10～10000 倍。但是，在使有这种方法前，要首先改变基因使其带有识别位点。由于需要预先的修饰，因此稀有位点的核酸内切酶如 I-*SceI* 不能用作常规工具来刺激基因打靶。也不能应用限制酶，因为它的酶切位点太普遍了。以一个专有的 DNA 结合区与非专一性切割位点相结合的嵌合性核酸酶，在原理上可以提供给基因组一个专有位点的双链断裂。

② 形成三链体的寡聚核苷酸（triplex-forming oligonucleotides，TFOS）　TFOS 与双螺旋 DNA 上的多聚嘌呤位点有很高的亲和性，这样就形成三螺旋 DNA 结构。TFOS 结合的专一性来自 TFOS 和目标双螺旋中富含嘌呤的链的大沟之间形成的氢键。TFOS 能通过向基

因组上的靶位提呈一种化学性损伤介质，来引入位点专一性损伤。补骨脂素常与 TFOS 一起使用，因为它稳定、无毒，当暴露在长波紫外光线下时（不能伤害细胞）能有效地形成单加成物和链内交联。三链体诱导的 DNA 损伤能刺激同源重组，但程度差异很大，这种差异性是由几种因素引起的，包括细胞类型、三链体的靶位点、损伤的类型、TFOS 在细胞中的稳定性等。其中，TFOS 的吸收和稳定性可能是限制性因素。

4. 基因打靶技术的应用

基因打靶的应用研究取得了丰硕的成果。基因打靶技术与小鼠胚胎干细胞技术结合，产生了近千种转基因小鼠品系，每个品系都包含着精确的基因修饰，包括基因敲除、基因敲入和精巧的突变。对于大动物的胚胎干细胞还没有成功的报道，但已成功地用培养的体细胞进行基因打靶。PPL 公司的研究人员在原代绵羊和猪细胞中进行了基因打靶，并通过体细胞核移植，获得了带有目的基因的克隆羊。基因打靶在植物上的应用研究才刚起步。Park JM 等人对拟南芥 ADL 1 基因进行了基因敲除。Ruf S 等对烟草进行基因打靶，发现了一个新的与光系统 I 相关的基因。Schaefer DG 等利用藓类植物（*Physcomitrella patens*）作为基因打靶模式植物进行实验研究也取得了很大的进展。George C. Allen 等利用 SARs（核骨架结合序列）构建基因打靶载体用于烟草的转化，提高了 *GUS* 基因的表达。周杰等利用基因打靶技术将人肝金属硫蛋白突变体 β 基因整合到 *Synechocystis* sp. PCC6803 的染色体上，并在蓝藻中得以表达。基因打靶技术在微生物领域有广泛的应用。在 20 世纪 80 年代，利用酵母 35S rDNA 作为同源序列将外源基因插入到酵母 rDNA 区，使外源基因得到稳定表达。在随后的研究中，人们利用酵母的 rDNA 作为同源序列构建了一系列高转化效率、高拷贝插入及高度表达的打靶载体。在寄生虫学的研究中也应用了基因打靶技术，利用此技术制备的利什曼原虫的减毒活虫疫苗已进入临床前期研究。

（三）时空可调节性基因打靶技术

前文所述几种打靶策略均为完全基因敲除，无论是置换还是插入策略，所导入的外源筛选基因片段都可能对基因组造成不可恢复的干扰，从而影响基因组 DNA 功能的变化。许多在成体器官发育中有重要功能的基因，由于在胚胎早期就被敲除，导致胚胎早期死亡，无法对该基因进行深入研究，为此 Cre-loxP 和 FLP-frt 等条件性敲除系统应运而生，使得组织特异性基因敲除变为可行，并最终实现时空可调节的打靶。所谓条件性基因打靶（conditional gene targeting）可定义为将某个基因的修饰限制于小鼠某些特定类型的细胞或发育的某一特定阶段的一种特殊的基因打靶方法。它实际上是在常规的基因打靶的基础上，利用重组酶介导的位点特异性重组技术，在对小鼠基因修饰的时空范围上设置一个可调控的“按钮”，从而使对小鼠基因组的修饰的范围和时间处于一种可控状态。利用 Cre/LoxP 和来自酵母的 FLP-frt 系统可以研究特定组织器官或特定细胞中靶基因灭活所导致的表型。

Cre/LoxP 系统是由 Hua Gu 和 JD Marth 等人在 1993 年首先提出的。它包括 Cre 重组酶和 *loxP* 位点两部分。前者为来自 *E. coli* 噬菌体 P1 的 *cre* 基因编码，*loxP* 由两个 13bp 的反向重复顺序和 8bp 的间隔区域构成。Cre 重组酶可介导 34bp 的重复单元，切除同向重复的两个 *loxP* 位点间的 DNA 片段和一个 *loxP* 位点，保留一个 *loxP* 位点。简单地说，Cre 是位点特异性重组酶，一定条件下可以切割 *loxP* 位点，催化两个 *loxP* 位点发生重组，以达到使 *loxP* 位点之间的基因缺失的目的。

下面是 *loxP* 位点的核苷酸组成。

ATAACTTCGTATA　ATGTATGC　TATACGAAGTTAT

反向重复序列　间隔区　反向重复序列

通过常规基因打靶在基因组的靶位点上装上两个同向排列的 *loxP*，并以此两侧装接上 *loxP* 的（“loxP floxed”）ES 细胞产生“loxP floxed”小鼠，然后，通过将“loxP floxed”小鼠与 Cre 转基因鼠杂交（也可以其他方式向小鼠中引入 Cre 重组酶），产生靶基因发生特定方式（如特定的组织特异性）修饰的条件性突变小鼠。在“loxP floxed”小鼠，虽然靶基因的两侧已各装上了一个 *loxP*，但靶基因并没有发生其他的变化，故“loxP noxed”小鼠表型仍同野生型一样。但当它与 Cre 转基因小鼠杂交时，产生的子代中将同时带有“loxP floxed”靶基因和 Cre 基因。Cre 基因表达产生的 Cre 重组酶就会介导靶基因两侧的 1*oxP* 间发生切除反应，结果将一个 *loxP* 和靶基因切除。这样，靶基因的修饰（切除）是以 Cre 的表达为前提的。Cre 的表达特性决定了靶基因的修饰（切除）持性，即 Cre 在哪一种组织细胞中表达，靶基因的修饰（切除）就发生在哪种组织细胞；而 Cre 的表达水平将影响靶基因在此种组织细胞中进行修饰的效率。所以，只要控制 Cre 的表达特异性和表达水平就可实现对小鼠中靶基因修饰的特异性和程度。具体打靶策略见图 11-9。

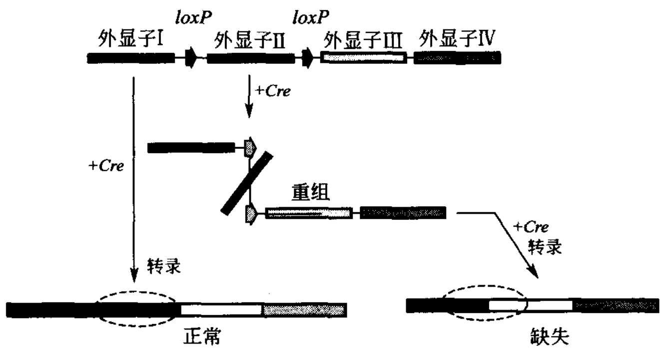


图 11-9　Cre-loxP 打靶策略示意图

有两种操作方法应用 Cre/loxP 系统：在构建打靶载体时，将标记基因放在靶基因内部，标记基因的两侧放上相同方向排列的 *loxP* 序列，而后既可以在细胞水平上用 Cre 重组酶表达质粒转染中靶细胞，通过识别 *loxP* 位点将抗性标记基因切除；或者在个体水平上将打靶杂合子小鼠与 Cre 转基因小鼠杂交，筛选子代小鼠就可得到删除外源标记基因的条件性敲除小鼠。Cre 转基因小鼠是将 *cre* 基因与各种组织（位点、时间、发育阶段）特异性的启动子连接构建载体，用传统的转基因技术或基因敲除技术获得相应的转基因小鼠。当 Cre 重组酶基因与可诱导的启动子连接时，即可通过诱导表达 Cre 重组酶而将 *loxP* 位点间的基因切除，从而实现特定基因在特定时间的失活。然而，现阶段可利用的组织特异性表达 Cre 重组酶的转基因小鼠还很有限，Cre/loxP 系统的应用还将依赖于更多组织特异性标志基因的发现以及人工调控基因表达系统的进一步研究。

（四）基因陷阱

随着许多生物（如拟南芥、水稻、酵母）整个基因组的破译，尤其是人类基因组框架图的完成，有关基因组的研究已从结构基因组时代进入功能基因组时代，即需要了解究竟有哪些碱基对应着生物体的何种功能。功能基因组研究的主要内容是利用结构基因组所提供的信息，发展和应用新的实验手段系统地分析基因的功能。目前，鉴定功能基

因的主要方法有基因表达系列分析（SAGE）、cDNA微阵列、蛋白组技术、基因芯片、基于转座子标签和T-DNA标签的反向遗传学技术等。下面介绍一种新技术——基因陷阱（gene trap），它主要依靠报告基因的随机插入来产生融合转录物或融合蛋白，通过检测报告基因而推知基因及其功能。因这系列方法酷似以报告基因为诱饵来捕获基因，故得名基因陷阱。

1. 基因陷阱的原理

目前已发展了三种不同的陷阱系统：增强子陷阱，启动子陷阱和基因陷阱。它们之间最大的不同体现在报告基因重组体的组成和插入位置上。基因陷阱是这三种方法的统称。

（1）增强子陷阱（enhancer trap）　将某报告基因与一个精巧的启动子相连，组成一增强子陷阱重组体（见图11-10），它不会自主起始转录，而需由被插入的细胞基因组中的增强子帮助才可转录。若报告基因最终表达，则可推知插入位点附近有增强子，或有基因，即实现了以该增强子陷阱重组体发现增强子的目的。

（2）启动子陷阱（promoter trap）　通过将报告基因插入到细胞基因组的外显子上［见图11-10(c)］，一旦发现它与细胞基因组基因被共同转录或表达，则可知该报告基因附近有启动子，从而起到了以之为诱饵，发现启动子的目的，这就是启动子陷阱。

（3）基因陷阱（gene trap）　基因陷阱重组体由报告基因和接受体剪接部位（splice acceptor，SA）组成（接受体剪接部位在报告基因上游），该重组体需要插入到细胞基因组的内含子中随着基因转录和表达［见图11-10(d)］，故如能检测到融合转录物或融合蛋白，则证明插入位置附近有基因存在。

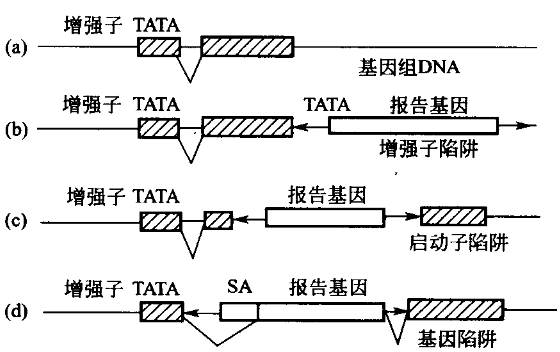


图11-10　各种基因陷阱的结构示意图

(a) 一条基因组DNA的某个基因（外显子为带斜线盒状部分；内含子为线状且用V形注出，表示会被切除；各基因陷阱重组体为空白盒状）。(b) 增强子陷阱。由启动子（TATA）和报告基因构成的重组体，插入到基因的增强子附近（图中所示为第二个外显子下游），报告基因的启动子被该基因的增强子激活，导致报告基因表达 (b)、(c)、(d) 中带黑箭头区域表示外源插入部分）。(c) 启动子陷阱。报告基因插入到基因外显子（图中所示为第二个外显子）之中，并借助该基因的增强子和启动子被转录和表达。(d) 基因陷阱。受体剪接部位和报告基因重组体插入基因内含子中，则转录时受体剪接部位与该基因内含子一起被切除，得到融合基因转录物

启动子陷阱和基因陷阱都有位置限制，前者需插入到外显子，后者则需插入到内含子，增强子陷阱却没有此位置限制；由于增强子与基因的位置可近可远，故增强子陷阱不易定位基因；此外，对启动子陷阱和基因陷阱而言，插入可能导致基因失活。

2. 基因陷阱的特点

基因陷阱的优势在于它只在表达水平上定位基因，细胞基因本身的转录和表达不受影响，所以可检测功能上多余的基因，也可检测在基因表达多个水平上都有作用的基因，或低水平表达的重要基因，而以前的功能基因组学的方法对这些基因都是无法确定的。

基因陷阱的不足主要表现在以下几个方面。

（1）表达基因的确定较困难且耗时长　如何克隆和筛选融合转录物或融合蛋白是一件很费时费力的事。通常根据所插入的报告基因的 DNA 序列为引物，用 PCR 技术获取两侧序列，再与 EST 数据库比较和确定基因功能。最近出现一些很好的解决方法，如在 RACE 的微透析和固相测序中采用自动化和荧光测序，可缩短时间；或只测一小段 DNA 序列即与 EST 数据库已有序列进行比较，若已有人研究过，则不用继续，若从没研究过，则应得到目的基因全长并分析其功能。

（2）突变的细胞的表型不明显　主要原因是被捕获的基因不一定有很明显的表达方式，以及突变的 ES 细胞克隆不一定有明显的表型。目前的解决方法是进行体外预先筛选我们需要的融合基因株系。

（3）报告基因表型可能会丧失　若报告基因没按预期加以整合，或报告基因被剪接时整个被切除，都可能使报告基因表型丧失而无从着手。

（4）目前使用的基因陷阱系统对细胞的损伤大　毕竟基因陷阱是插入了一段外源 DNA，所以是一种剧烈的方法，会有致死效应。在一些情况下应采用柔和一些的方法，如进行重组体的置换而不是插入、用点突变、选择性地采用表型突变。

3. 基因陷阱的应用

基因陷阱的应用主要集中在医疗和制药上，而目前小鼠是探讨人类生理过程最好的模型，因此许多应用性的工作都以小鼠作为研究对象。一般在小鼠中不使用增强子陷阱，小鼠基因组很大但基因频率很低，故很难定位增强子；报告基因一般是 *lacZ*，它的上游有一个或多个受体剪接部位，由逆转录病毒引入全能胚胎干细胞，再移入小鼠种系中。1988 年就已有人在小鼠细胞中使用基因陷阱。最近美国国立卫生院（NIH）的科研人员报告以小鼠为材料，利用 HIV1 慢病毒来研究哺乳动物的细胞分化和发现新的基因，而慢病毒作为基因陷阱载体，是因为它们可随机插入动物细胞的基因组并可长期表达。法国和日本的科学家利用乙肝病毒（HBV）的 DNA 帮助基因陷阱的插入整合，来确定未知的肝癌相关基因。这说明可通过研究病毒基因的整合位点来确定人类疾病相关基因，同时证明这些由 HBV 构成的基因陷阱附近的基因对细胞增殖和生长有重要调控作用。近些年来，许多著名的大学和研究所已建立自己的基因陷阱实验室，如英国牛津大学、美国冷泉港实验室等。此外，许多生物制药公司和研究团体都看好基因陷阱在生物医疗和制药方面的巨大潜力，如美国 Lexicon Genetics Inc. 以鼠为材料探讨疾病的分子基础，研究分泌蛋白、蛋白偶联受体、离子信道蛋白和转运蛋白等来用于医药开发，为此而建立了自己的基因陷阱数据库和小鼠克隆总的文库，从而为生产和应用创造了一个整合的功能基因组平台。Bay Genomics 在鼠胚胎干细胞中利用基因陷阱来失活一些基因，而这些失活某些基因的干细胞可产生类似基因敲除的小鼠个体，利用这些小鼠研究与动脉粥样硬化和脂肪代谢有关的基因，探索气喘、肺结核的病理学，同时结合 DNA 微阵列和原位杂交来研究与心肺发育和致病有关的基因。

（五）诱变技术在功能基因组学中的应用

基因打靶可使许多等位基因发生突变，但由于基因打靶的特点，这些突变常常集中在特定的基序（motif）或结构域（domain）。然而在一个基因中，有时一段很短的片段，甚至一个碱基的改变就可以造成功能的变化。随着人类及小鼠基因组测序工作的完成，我们需要更详尽地了解基因各个部分的功能，诱变技术的优势便显现出来了。

1. 定向诱变

定向诱变（targeted mutagenesis）是利用同源重组技术，使胚胎干细胞（embryonic stem cell，ES cell）内目的基因产生定点突变。这些突变可进一步用于基因敲除、转基因动物、显性负突变等研究。最近两年发展了许多构建靶结构的新方法，如酿酒酵母中微同源重组（microhomologous recombination），通过 PCR 方法产生一个特定的靶 DNA 片段，这个片段含有一个两侧带有与酵母基因同源的 35～50bp 作为选择性标记，就足以促进酵母的同源重组，而在小鼠 ES 细胞中，至少需要 119kb 的连续基因组 DNA 才能产生有效的同源重组。以往这种方法只用在酵母中，现在也用到小鼠上。

2. 表型诱变

定向诱变方法是用于已知基因的突变，而表型诱变（phenotype-driven mutagenesis）是用于未知基因的诱变。表型诱变的主要优点是无需知道哪个基因以及这些基因的何种突变导致特定的表型或疾病。用表型诱变剂进行诱变后，可以用筛查整个基因组的办法来寻找新的显性或隐性突变。该方法需要大量的小鼠杂交群体，工作量较大，但这种全基因组扫描法是筛查整个基因组中单一突变的最好方法，因为任何一个导致一定表型的可能突变都可以被检测出来。

人类基因组和模式生物体基因组全序列分析，可为用比较基因组序列的方法来研究基因的功能提供极大的可能。大量的模式生物对于人类基因组和基因组序列的研究，为深入研究基因功能提供了非常好的机会。模式生物对于医学的研究也将起到重要的作用。果蝇在遗传和 X 射线诱导突变的染色体原理的早期研究，促进了遗传学的发展；线虫、斑马鱼、其他小生物的研究现已进入新的发展时期。通过在模式生物中构建与人类疾病相关的特殊细胞通路模型，将发展人类疾病全新治疗方案。人类疾病基因在模式生物中研究的突破，无疑将给疾病的预防、诊断和治疗带来深刻的革命。

七、SNP、EST 在研究新（未知）基因中的应用

（一）单核苷酸多态性（SNP）

人类基因组中存在广泛的多态性，最简单的多态形式是发生在基因组中的单个核苷酸的替代，即单核苷酸多态性。它是继第一代限制性片段长度多态性（RFLP）标记、第二代微卫星（MS）即简单的串联重复标记后，出现的第三代基因遗传标记。SNP 是指基因组水平上单个核苷酸变异引起的 DNA 序列多态性，单碱基的变异主要由转换（transition）或颠换（transversion）所引起。SNP 一般只有 2 个等位基因和 3 种基因型，又称为双等位基因标记（biallelic marker）。SNP 遍布于整个人类基因组中，根据 SNP 在基因中的位置，又可分为基因编码区 SNP（coding-region SNP，cSNP）、基因周边 SNP（perigenic SNP，pSNP）以及基因间 SNP（intergenic SNP，iSNP）等 3 类。

1. SNP 检测方法

SNP 表现为单个碱基的变异，理论上讲任何能检测点突变的方法均可用于 SNP 的识别

和检出，主要的技术方法包括单链构象多态性法（single strand conformation polymorphisms，SSCP）、异源双链分析（heteroduplex analysis，HA）、高效液相层析法（HPLC）、DNA 直接测序分析、动态特异等位基因杂交（dynamic allele-specific hybridization）、变异检测阵列法（variant detector arrays，VDA）、基因芯片、TaqMan 技术以及基质辅助激光解吸附电离飞行时间（MALDI-TOF）质谱法等。目前，许多生物技术公司发展出高通量检测 SNP 的技术系统，如荧光微阵列系统（Affymetrix）、荧光磁珠技术（Luminex，Illumina，Q-dot）、自动酶联免疫（ELISA）试验（Orchid Biocomputer）、焦磷酸的荧光检测（Pyrosequencing）、荧光共振能量转移（FRET）（Third Wave Technologies）以及质谱检测技术（Rapigene，Sequenom）等。

2. SNP 网上资源

由于人类基因组工程的带动，许多物种都已开始了基因组的项目，并建立了大量的数据库，比较这些来源于不同实验室不同个体的序列，就可检测到 SNP，Internet 的使用，使这项工作变得十分简单，国内外许多学者已着手这方面的工作，并开发了一些应用软件。

以下是可供利用的公开 SNP 网上资源。

① 由美国国立卫生研究院（National Institutes of Health，NIH）提供的主要是与癌症和肿瘤相关的候选 SNP 数据库：http：//cgap. nci. nih. gov/GAI。

② 由 NIH 开辟的适于生物医学研究的 dbSNP 多态数据库：http：//www. ncbi. nlm. nih. gov/SNP。

③ 由人类基因组组织机构（Human Genome Organization，HUGO）维持的突变数据库：http：//ariel. ucs. unimelb. edu. au/～cotton/mdi. htm。

④ 由美国 Whitehead 研究所（Whitehead Institute for Biomedical Research Genome Institute）建立的人类 SNP 数据库：http：//www-genome. wi. mit. edu/SNP/human/index. html。

⑤ 由华盛顿大学（Washington University）支助的按染色体位置组织的 SNP 数据库：http：//www. ibc. wustl. edu/SNP。

⑥ 由瑞典卡尔林斯卡研究院（Karolinska Institute of Sweden）建立的 HGBase 数据库：http：//hgbase. cgr. ki. se/。

⑦ 由国际医药与信息加工公司联合组成的 SNP 研究联盟（The SNP Consortium，TSC）建立的 SNP 数据库：http：//snp. cshl. org/db/snp/map。

⑧ 由美国国立环境健康科学研究院（National Institute of Environmental Health Science）资助的犹他州大学 SNP 数据库：http：//www. genome. utah. edu/genesnps/。

3. SNP 的应用

与前两代分子标记相比，SNP 具有以下特点及优点：①数量多、覆盖密度大，据估计平均每 1000bp 就会出现一个。目前已检测出 93%的人类基因包含 SNP，98%的 SNP 其距离不超过 5kb，因此可以在几乎每个基因的内部或附近提供一系列标记；②由于 SNP 一般只有 2 个等位基因，在检测时只需要通过一个简单的“+/—”方式即可进行基因分型，这使得检测、分析易于实现自动化；③遗传稳定性强，与微卫星等重复序列多态性标记相比，SNP 具有更高的遗传稳定性；④多态性丰富，SNP 包含了目前已知 DNA 多态性的 80%以上，是最常见的遗传变异类型；⑤易于基因分型。SNP 的二态性，也有利于对其进行基因分型。对 SNP 进行基因分型包括 3 方面的内容：①鉴别基因型所采用的化学反应，常用的

技术手段包括DNA分子杂交、引物延伸、等位基因特异的寡核苷酸连接反应、侧翼探针切割反应以及基于这些方法的变通技术；②完成这些化学反应所采用的模式，包括液相反应、固相支持物上进行的反应以及二者皆有的反应；③化学反应结束后，需要应用生物技术系统检测反应结果。

SNP的双等位基因及分布广泛的特点适于自动化大规模扫描，成为继微卫星之后最受推崇的作图标记。国际SNP作图工作组2001年发表的人类基因组图已定位了142万个SNP，平均间距达到1191kb。SNP被认为是一种稳定遗传的早期突变，其标记间存在的连锁不平衡与疾病的相关性较为稳定，可作为识别、定位疾病相关基因的有力工具；另外，存在于编码区的SNP可能改变蛋白质产物的性质及功能，而编码区及其邻近调控序列的SNP可影响基因的表达产物和表达水平。筛选并研究这些标记将对疾病诊断、防治和药物设计起到积极作用；同时将有助于对基因的作用机理做深入分析，进而探明复杂性状内参与作用的微小效应。目前SNP研究主要包括2个方面：一是确定人类究竟有多少SNP位点；二是直接利用一些患病家系，选择与致病有关的候选基因上的SNP做连锁分析。研究SNP的意义在于能更好解释个体的表型差异，不同群体或个体对肿瘤的易感性、对药物的耐受性等问题，可广泛应用于致病基因定位、疾病的连锁分析、易感基因的检测等。

（二）表达序列标签（EST）

EST是来自随机选取的cDNA克隆的末端序列，即一个EST就是对应于某一种mRNA的一个cDNA克隆的一段序列。如前所述，在通过mRNA差异显示、代表性差异分析等方法获得未知基因的cDNA部分序列后，研究者都迫切希望克隆到其全长cDNA序列，以便对该基因的功能进行研究。克隆全长cDNA序列的传统途径是采用原位杂交的方法筛选cD-NA文库，或采用PCR的方法。这些方法由于工作量大、耗时、耗材等缺点已满足不了人类基因组时代迅猛发展的要求。在基因结构、定位、表达和功能研究等方面都积累了大量的数据后，如何充分利用这些已有的数据资源，加速人类基因克隆研究，同时避免重复工作，节省开支，已成为一个急迫而富有挑战性的课题摆在我们面前，采用生物信息学方法延伸EST序列，获得基因部分乃至全长cDNA序列，将为基因克隆和表达分析提供空前的动力，并为生物信息学功能的充分发挥提供广阔的空间。

1. EST与基因识别

EST技术最常见的用途是基因识别，传统的全基因组测序并不是发现基因最有效率的方法，这一方法显得既昂贵又费时。一部分科学家支持首先对基因的转录产物进行大规模测序，即从真正编码蛋白质的mRNA出发，构建各种cDNA文库，并对库中的克隆进行大规模测序。虽然EST序列数据相对不精确，精确度最高为97%，但实践证明EST技术可大大加速新基因的发现与研究。应用EST技术，可以跳过生物分类学的界限，从生物模型的已识别基因迅速克隆出人和小鼠基因组中相应的更复杂的未知基因。生物间在核苷酸水平上的进化差异阻碍了传统意义上的杂交或以PCR为基础的基因克隆策略，即使是亲缘关系很接近的生物也不例外。通过计算机进行dbEST进行数据库筛选，其本质是电子杂交实验，提供了一条更为广泛的基因识别路线。

2. EST与“电子”基因克隆

利用计算机来协助克隆基因，称为“电子”基因克隆（sillcon cloning），是与定位克隆、定位候选克隆策略并列的方法之一，它采用生物信息学的方法延伸EST序列，以获得基因部分乃至全长的cDNA序列。

基本原理及步骤如下。利用计算机协助克隆基因的第一步是必须获得感兴趣的 EST，在 dbEST 数据库中找出 EST 的最有效途径是寻找同源序列（标准：长度≥100bp，50%≤同源性 50%≤85%），可通过 BLAST 检索程序实现，然后将检出序列组装为重叠群（contig），以此重叠群为被检序列，重复进行 BLAST 检索与序列组装，延伸重叠群系列，重复以上过程，直至没有更多的重叠 EST 检出或者说重叠群序列不能继续延伸，有时可获得全长的基因编码序列，再与 GenBank 核酸数据库进行相似性检测。假如没有精确匹配基因，将 EST 序列数据按 EST 六种阅读框翻译成蛋白质，接着与蛋白质序列数据库进行比较分析。“电子”克隆新基因的基本流程见图 11-11。

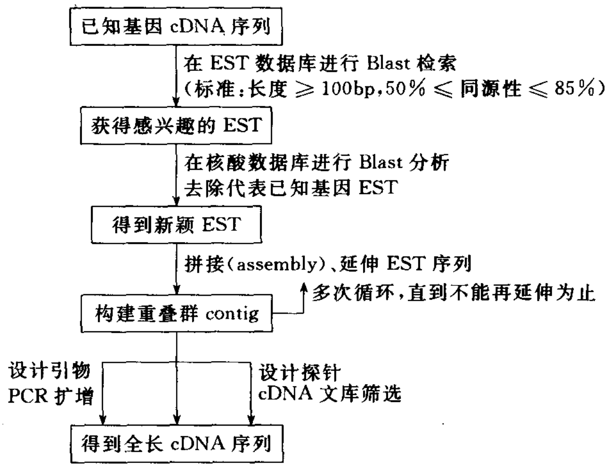


图 11-11 “电子”克隆新基因的基本流程

然后利用 UniGene 数据库进行电子延伸，即获得待分析序列的 UniGene 编号以后，将参与形成 UniGene Cluster 的所有核酸序列下载到本地，利用 Sequencher Tm 软件（Mac 版）或其他序列装配软件进行组装，形成较长的新生序列。同时，进行开放阅读框（open reading frame，ORF）分析和含完整编码区的 cDNA 的克隆及验证。

3. 电子表达谱分析与基因的电子定位

利用 UniGene 数据库以及利用 Tigem 的电子原位杂交服务器进行电子表达谱分析；对核酸序列进行电子基因定位，从而了解其在染色体上的定位信息。进行电子基因定位一般有两种策略：①通过序列标签位点（sequence-tagged site，STS）进行定位（电子 PCR 程序）；②通过 UniGene/RH 技术进行定位。

（伍欣星）

第十二章　蛋白质组学技术

蛋白质组的英文是 PROTEOME，它来源于 PROTEins 和 genOME 两个词的组合，意思是 proteins expressed by a genome，即由基因组表达的蛋白质。1994 年这一概念被首次提出后，随之兴起的是一个新的学科领域——蛋白质组学（proteomics）。从广义上来讲，蛋白质组是指某种细胞或组织中基因组表达的“所有蛋白质”。与基因组的概念不同之处在于：蛋白质组是一个动态的概念，“所有蛋白质”其实是一个模糊的说法，因为即使是同一种细胞，处于不同的状态，如发育的不同阶段、外界的不同影响等，细胞中的蛋白质也会有相应的变化，从而导致不同的蛋白质组。因此，有人提出了功能蛋白质组（functional proteome）的概念，指细胞在一定阶段或与某一生理现象相关的所有蛋白质，这实际上是把蛋白质组的概念具体化。从狭义上来讲，蛋白质组可以指不同时期细胞内蛋白质的变化，比如正常细胞和异常细胞之间、细胞用药和不用药之间的蛋白质水平上的差别，这在疾病研究和药物筛选上很有意义，也是蛋白质组在应用上最具前景的方面。当然，随着蛋白质组领域研究的不断发展，蛋白质组的概念也将不断延伸、深入和分化。

第一节　概　　述

一、蛋白质组学研究的意义和背景

如何了解生命现象的过程及其变化，如何解释疾病的发生，如何诊断和治疗疾病，一直是生命科学的核心。核酸和蛋白质是控制生命过程的两种主要大分子，因此也是研究的主要对象。传统的蛋白质研究是针对单个蛋白质的，但由于蛋白质的可变性和多样性，对蛋白质的分离分析是较困难的。因此，当人们意识到基因携带着生命功能的基本信息，并可以一定程度上控制蛋白质的结构与功能时，基因受到了更多的重视，加上技术上越来越成熟，使基因研究日益广泛。但是单个基因的研究还不能反映生命现象的整个过程，因此出现了大规模的基因组研究，以期在整体水平上了解基因的活动规律。但基因组计划完成以后，是否就能解决生命现象中的所有问题呢？目前有二十几个物种已完成了基因组的全序列测定，但在这些物种的基因组中，至少有 40%～60%的基因编码的蛋白质功能是未知的。目前，功能基因组中所采用的策略如 cDNA 阵列（cDNA microarray）、DNA 芯片（DNA chips）、基因表达序列分析（serial analysis of gene expression，SAGE）、表达序列标记（expressed sequence tag，EST）等都是从细胞中 mRNA 的角度来考虑的，而这一考虑的前提是细胞中 mRNA 的水平反映了蛋白质表达的水平。但事实并不完全如此，从 DNA→mRNA→蛋白质，存在 3 个层次的调控，即转录水平调控、翻译水平调控、翻译后水平调控。以 mRNA 角度考虑，实际上仅包括了转录水平调控，并不能全面代表蛋白质表达水平。实验也证明，

组织中 mRNA 的丰度与蛋白质的丰度并不完全一致。更重要的是，蛋白质复杂的翻译后修饰则几乎无法从 mRNA 水平来判断。因此，很多分子生物学家认识到，即使是完成了基因组计划，也不能完全就此解决预期的问题。作为生命功能的行使者，蛋白质的重要作用不言而喻，它在某种程度上比基因更能直接地反映生理功能的过程及其变化。蛋白质本身的存在形式和活动规律如翻译后修饰、蛋白质间相互作用以及蛋白质构象等问题，仍依赖于对蛋白质的研究来解决。虽然蛋白质的可变性和多样性一定程度上是技术上难以把握的问题，但正是这些特性参与和影响着整个生命过程，是生命现象复杂性的真实体现。

传统意义上的蛋白质研究，往往以单个或某一类蛋白质作为对象，显然不能反映生命现象的复杂性和可变性，也不能与基因组这样整体水平的研究相匹配。基因组计划的快速推进，实际上给蛋白质组的研究提出了挑战和机遇。一方面，需要将单一对象的研究转变到细胞乃至组织的整体蛋白质研究，这不仅在理论上是一个全新的思路，更在实践中对蛋白质组研究技术体系提出了挑战；另一方面，基因组计划的进行，越来越多物种基因全序列的获得，大量 EST 的确定，为蛋白质的快速鉴定建立了很好的基础。在这样的条件和背景下，蛋白质组研究已经成为一个新的研究热点，这一领域的发展将大大加速生命科学的研究进程，并可望在疾病研究方面起到重要作用。

二、蛋白质组学研究的特点

随着“后基因组时代”的到来，功能基因组成为研究的热点，而蛋白质组学研究则是功能基因组研究的“中流砥柱”，是这一前沿研究的战略制高点。相对于基因组的研究，蛋白质组学的研究有自身的规律和特点。

1. 同一性与多样性

基因组作为遗传信息的载体，其最主要的特征就是同一性。对于单细胞生物而言，不论在什么样的生长条件下，其基因组始终是不变的。对于多细胞生物来说，同一个体的基因组不论是在不同的发育阶段或不同种类的细胞里都是同样的。知道了个体内某一细胞内的基因组就等于知道了该个体所有细胞的基因组。然而，对于蛋白质组而言，由于蛋白质是生命活动的主要执行者，因此，对于不同类型的细胞或同一个细胞在不同的活动状态下，蛋白质组的构成是不一样的。

任何一种基因组都是由不编码蛋白质的核苷酸序列和编码蛋白质的核苷酸序列——基因所组成。然而，除了已经知道有蛋白质产物的基因以外，其余的“基因”只不过是根据“起始密码”和“终止密码”序列所确定的“开放阅读框”（open reading frame，ORF）。一般说来，一个基因组内的基因数目通常是指 ORF 的数目。基因组研究的实践告诉我们，要想准确无误的肯定一个基因组内的 ORF 数目，并不是一件容易的事。虽然生物信息学已经有了长足的进步，但是直接从基因组的序列进行基因注释（annotation）还是很困难的。不久前人们发现，对一个简单的细菌（*Mycoplasma genitalium*）基因组的 340 个基因的注释，其错误注释的比例至少是 8%。对于那些高等生物复杂的基因组，尤其是存在 RNA 剪接的基因组，要想确定其 ORF 的数目就更成问题了。

显然，最能确定 ORF 的数量应该是通过研究蛋白质组来进行，因为确定了一个蛋白质就肯定了一个基因，但确定了一个 ORF 却并不能肯定它会有对应的蛋白质产物。据最新统计，人类基因组拥有的基因数目为 3 万～4 万个。如果我们能够把人的 252 种细胞内的全部蛋白质都鉴定出来，那么就有可能真正知道人类基因组的所有基因。

2. 有限与无限

对于基因组研究而言，核心任务是测定所有的 DNA 碱基序列。例如，芽殖酵母基因组有 1200 万碱基对，人类基因组的最新统计为 32 亿个碱基对。要测定的基因组不论大小，其核苷酸的数量是明确的。然而，对蛋白质组来说，蛋白质的种类究竟有多少就很难说了。上面说过，蛋白质组可以被定义为基因组基因表达的所有蛋白质，但这一定义没有考虑蛋白质的修饰。细胞内的大部分蛋白质通常都被进行过翻译后修饰，主要有磷酸化、糖基化、酰基化等。修饰过的蛋白质具有特定的物理化学性质和生物学功能。对蛋白质修饰的研究已构成了蛋白质组研究的一个重要部分，由此还形成了蛋白质组学的一个分支——修饰蛋白质组学。如果把一个修饰蛋白视为一种新的蛋白质，那么蛋白质组的蛋白质数量将远远大于相应的基因组的基因数量。在这个意义上，估计人类蛋白质组的蛋白质的种类为 20 万～200 万。显而易见，对蛋白质组的蛋白质数量的估计是非常模糊的。

从蛋白质修饰的角度来看，不仅仅是蛋白质组的蛋白质种类大大增加，更重要的是由于不存在度量蛋白质修饰的尺度，人们也许永远不能像确定基因组核苷酸序列那样肯定地把握生物体内蛋白质组的蛋白质数量。如果说表达的蛋白质的种类可以根据基因的数目来确定，那么修饰形成的蛋白质种类只有依靠对蛋白质的直接研究来确定，而这种研究是没有止境的，因为蛋白质的修饰和生命的活动密切相关。

3. 静态与动态

一个个体的基因组自个体诞生到死亡，始终保持不变；而作为新陈代谢的主要执行者的蛋白质组，在个体的生命活动中却总是变动不停。生命活动中的蛋白质可以大致分为两类：一类是相对比较稳定的，如负责细胞各种结构组成的蛋白质；另一类则是随细胞的活动和状态发生量或质变化的蛋白质，如负责信号转导或细胞活动调控的蛋白质。因此，人们可以通过确定细胞中蛋白质的变化来解释细胞的生物学功能的变化。蛋白质组研究的一个重要任务是找出蛋白质组里发生了变化的蛋白质。人们把以研究蛋白质变化为主的蛋白质组学称为功能蛋白质组学（functional proteomics）或比较蛋白质组学，它主要是比较在不同生长状态下或病理情况下的同一种细胞或组织的蛋白质组内的差异蛋白质。与此相对应，测定一个细胞或组织含有的全部蛋白质种类的研究就是通常所说的蛋白质组学。

上述对蛋白质动态的研究又引出了蛋白质组学的另一个问题，即怎样区分差异蛋白质是属于生物学意义上的还是实验意义上的，因为蛋白质的稳定性要比核酸差得多，在制备的过程中可能会发生降解或丢失。此外，差异蛋白质的区分通常是通过蛋白质双向凝胶电泳来检测，而双向电泳由于自身分辨率的限制，难以测定微量的蛋白质变化。有实验表明，目前的双向电泳技术对酵母细胞中低于 1000 个拷贝量的蛋白质是检测不到的。还有一个问题是，对于大多数差异蛋白质而言，有或无这类质的变化的差异蛋白质较少，而在量的增减方面的差异蛋白质则较多。由于双向电泳技术的不稳定性，所以要准确测定蛋白质量的差异是一个困难的任务。因此，从动态的角度来研究蛋白质组是当前蛋白质组学的一个挑战。

尽管测定差异蛋白质的技术难度较大，但功能蛋白质组学在今后一段时间将仍然是蛋白质组研究领域的主要方向。首先，蛋白质的差异和变化是生命活动的重要方式：发现在正常和病理状态下出现差异的蛋白质常常是认识和防治疾病的关键。因此，大多数研究者仅仅关心寻找蛋白质的差异表达谱；其次，鉴定出一个细胞或组织中的所有蛋白质是一项费时费钱的工作，需要大量的人力和物力。相比之下，测定差异蛋白质的任务需要的投入就小得多。

4. 时间与空间

DNA 通常位于细胞核内，且保持稳定，因此测定基因组的 DNA 序列不受时空的影响。对于转录的 mRNA 来说，时间是主要的参考因素，在发育的不同阶段或细胞的不同活动时期，mRNA 的表达是不一样的。因此，在研究转录组或基因芯片时必须考虑到时间，但通常不需要考虑空间的影响。而在蛋白质组的研究中，不仅要考虑时间的因素，更要考虑空间的影响。首先，不同的蛋白质分布在细胞的不同部位。它们的功能与其空间定位密切相关。要想真正了解蛋白质的功能，必须要知道蛋白质所处的空间位置；其次，许多蛋白质在细胞里不是静止不动的，它们在细胞里常常通过在不同的亚细胞环境里的运动发挥作用。例如，在细胞周期的调控过程中，磷酸酯酶 Cdc25 在细胞核内对蛋白激酶 Cdc2 进行去磷酸化，从而激活 Cdc2，使细胞进入分裂期。如果 DNA 损伤发生，Cdc25 会被一种 14-3-3 蛋白从细胞核内转运到胞质内，导致核内的 Cdc2 保持磷酸化和失活状态，使细胞停滞在 G2 期。细胞的信号传导和转录调控也常依赖于蛋白质在空间位置的变化和运动。因此，蛋白质组学中又派生了一个与空间紧密相关的新研究领域——亚细胞蛋白质组学，这种亚细胞蛋白质组可能是细胞器蛋白质组，如高尔基体蛋白质组；也可能是比细胞器还要小的组分——如核膜。对亚细胞蛋白质组的研究将有助于对亚细胞结构和功能的理解。此外，值得一提的是，分级分离技术对于研究细胞的整体蛋白质组也是很有用的。目前，双向电泳技术能分离的蛋白质数量最多是几千种，而一个细胞内的蛋白质数量通常会超过万种，尤其是高等真核生物的细胞，可能有数万种蛋白质。因此，用双向电泳技术分离全细胞的蛋白质组分的效果很不理想。如果利用细胞的分级分离技术，人们可以把细胞不同组分的蛋白质谱分别进行鉴定，然后再进行“组装”，从而得到细胞完整的蛋白质谱。

对于亚细胞蛋白质组学来说，关键技术是对细胞进行恰当的分级分离。由此，蛋白质组研究又面临着一个新的挑战：如何获得完整的亚细胞成分。在这种分级分离过程中，既要避免待研究的亚细胞组成蛋白质的丢失，又要避免非组成蛋白质的“污染”。另外，如何区别亚细胞组成蛋白质和由于生理作用而在细胞不同区域进行运动的蛋白质也是亚细胞蛋白质组的一个技术难点。例如，负责转运蛋白质的高尔基体，既有组成蛋白质，也有待输送的蛋白质，在研究高尔基体蛋白质组时应该加以区分。

5. 孤立行为与相互作用

基因组的基因表达的各种 mRNA 是彼此孤立的，互不干扰。但是 mRNA 的产物——蛋白质正好相反，它们彼此间有着广泛的相互作用。也许可以说，不存在不与其他蛋白质发生作用的“孤立蛋白质”。蛋白质的功能的实现离不开蛋白质与蛋白质或蛋白质与其他生物大分子之间的相互作用。所以，蛋白质组研究的一个主要内容是研究大规模的蛋白质的相互作用。科学家为此专门发明了一个新的名词——“相互作用组”(interactome)。

相互作用组研究可以分为两类。第一类是研究蛋白质相互作用的网络。细胞内的许多活动如信号转导等都是通过一个复杂而广泛的蛋白质相互作用网络来实现的。不久前，科学家通过大规模双杂交的手段，利用 27 种蛋白质对线虫的基因组进行了筛选，得到了参与线虫生殖器发育的 124 种发生相互作用的蛋白质，其中大约有 100 种是未知功能的蛋白质。除了大规模双杂交技术外，人们还使用了噬菌体显示技术（phage-display）和蛋白质芯片技术。对于大规模蛋白质相互作用分析而言，最突出的技术问题是虚假信号比较多，如双杂交技术容易出现假阳性，而噬菌体显示技术则容易产生假阴性。显然，如果把不同技术结合起来，可以把结果进行互补，从而更准确地描绘出蛋白质的相互作用网络。

相互作用组的另一类研究是蛋白质复合体组成的分析。蛋白质复合体通常可以分为两种。一种是结构型的蛋白质复合体，如核孔复合体；这一类通常比较稳定。另一种则是功能型蛋白质复合体，例如负责转录的转录蛋白复合体或者负责 DNA 复制的复制蛋白复合体；这类复合体只有在执行功能时才聚合在一起，因此很不稳定。对于结构型蛋白质复合体，通常采用分级分离的方法进行富集，然后再用经典的蛋白质组研究方法如电泳和质谱进行研究。但对于功能型蛋白质复合体，现有的方法就很难胜任了。最近，研究者发明了一种新的分离方法，可能有助于对这类蛋白质复合体的研究。首先是构造能够表达含有特定标签肽段（tag）的复合蛋白质的质粒，这些质粒表达的蛋白质在正常生理状态下与细胞内能够与其发生相互作用的其他种类蛋白质形成复合物；然后通过对标签肽段有专一亲和性的亲和分离柱进行提取，最后再通过电泳和质谱对这些蛋白质复合物的组分进行分析。发展对蛋白质复合体的研究技术将是今后蛋白质组学的一项重要任务。

6. 单一手段与多种技术

在基因组研究中，DNA 测序技术是一个最基本和最主要的工具，这是因为基因组的均一性和简单性使得一种单一的技术就能胜任基因组研究任务。但是，在蛋白质组研究中，需要的研究技术就远远不止一种，并且技术的难度也远远大于基因组研究技术。

蛋白质组研究技术可以简单地分为两大类。第一类是蛋白质组的分离技术，如双向电泳技术、细胞分级分离技术、亲和层析技术都是蛋白质组常用的分离技术。蛋白质组分离有以下难点：首先是量的问题，不同的蛋白质数量在细胞内可以有很大的差别，例如在酵母细胞里，有些细胞周期调控蛋白不到 100 个分子，而糖基酶可能有 200 万个分子。估计蛋白质量的动态差别可达到 10^6 数量级。其次是质的问题，蛋白质是由 20 种化学性质各异的氨基酸所组成，因此不同的蛋白质的物理化学性质差别很大。有些蛋白质易溶于极性溶剂，有些蛋白质则难溶于极性溶剂；有些蛋白质较稳定，有些蛋白质则易降解。因此，蛋白质的分离很难找到对所有蛋白质都通用的方法。

第二类是蛋白质组的鉴定技术，其核心是质谱技术。蛋白质芯片技术、噬菌体显示技术和大规模双杂交方法等也属于鉴定技术。现有鉴定技术面临的一个最大问题是，它们都依赖于已知的蛋白质或基因的序列作为检测的基础，通过比较来确定待测定的蛋白质。对于在已有数据库中没有记载的全新蛋白质进行“从头测定”（*denovo* identification）是一个很困难的任务，一般是利用 Edman 降解法进行 N 端序列测定，或化学降解法测定 C 端序列。

简而言之，蛋白质组研究对技术的依赖性和要求远远超过基因组学。目前，蛋白质组学的研究技术还有很多不完善之处，许多新技术正在研发之中。因此，蛋白质组学的发展是受技术限制的，也是受技术推动的。

7. 互补与互助

HGP 在基因组测序工作完成之后，面临的主要任务就是要进行基因注释（annotation），找出所有可能的基因或 ORF。一般情况下注释的出错率是比较高的，但是，如果有蛋白质组的数据作参考，就可以大大提高注释的正确率。此外，由于高等真核生物个体的基因大多是非连续性的，即由编码氨基酸的外显子和非编码氨基酸的内含子所组成，因此确定完整的 ORF 也非易事。显然，利用蛋白质的序列来推测和确定 ORF 的全长也是很有帮助的。

蛋白质组的许多工作也离不开对基因组的研究。这一点在蛋白质相互作用的研究中尤为突出。双杂交技术、噬菌体显示技术等就是以基因组数据和技术为基础的。在质谱测定蛋白

质的过程中，离不开对已知基因组的基因数据的比较。因此，蛋白质组和基因组是互相依赖、互为补充。

第二节　蛋白质组学研究的技术体系与路线

就技术方法而言，蛋白质研究技术比基因研究技术相对要复杂和困难，这是因为不仅氨基酸残基数量远多于核苷酸残基（20/4），而且蛋白质有着复杂的翻译后修饰如磷酸化、糖基化等，这给分离分析带来很多困难。蛋白质也不像基因一样能够体外扩增，因此在量上还受到限制。从技术来看，蛋白质组的研究实质上是在细胞水平上大规模地对蛋白质进行分离和分析，这与传统意义上的蛋白质研究是不同的，后者的研究对象通常是某一个或某一类蛋白质，虽然两者仅一字之差，但无论在概念上还是在技术方法上都有区别。蛋白质组的分离分析是在细胞水平上的，往往要处理上千种蛋白质，对蛋白质研究技术的灵敏度、精确度和分辨率有很高的要求，传统的蛋白质研究方法无法适应这种要求，因此新的技术方法的发展是必需的、也是必然的。从某种意义上来讲，正是由于一系列技术方法上的突破，特别是双向凝胶电泳和质谱技术的出现，使蛋白质组的研究成为可能。在分离技术方面，双向凝胶电泳是目前惟一最有效的方法，而质谱技术已成为最重要的分析方法，这两者是目前蛋白质组最主要的支撑技术。另外，生物信息学在蛋白质组的研究中也有举足轻重的作用。以下就这几方面进行简要介绍。

一、蛋白质组分离技术

如前所述，双向凝胶电泳是目前蛋白质组最有效的分离方法，是蛋白质组技术的核心。现在意义上的双向电泳，是指第一向以蛋白质电荷差异为基础进行分离的等电聚焦，第二向以蛋白质分子量差异为基础的 SDS-PAGE。由于这一方法具有较高的灵敏度和分辨率，它立即被应用在分离细胞内提取的复杂蛋白质混合物，还成功分离了大肠杆菌的超过一千种蛋白质。但它当时并没有得到更为广泛的应用，因为第一向在外加电场下形成 pH 梯度不够稳定，致使重复性差，即使是同一实验室的比较也较困难；另外，上样量少，通常要多次分离才能积累足够的量进行分析。在当时的条件下，蛋白质分离后的分析工作量十分庞大，几乎不可能完全鉴定分离到的上千种蛋白质。直到近年来，第一向 IEF 采用丙烯酰胺与不同 pH 的两性电解质形成固定的 pH 梯度，才改进了重复性和上样量这两个问题。同时，预制的商品化 IPG 胶条也使操作趋于自动化，加上分析鉴定技术的大幅度改进以及计算机数据库的应用，使双向凝胶电泳在分离细胞内蛋白质方面有其他方法所无法比拟的优势，成为蛋白质组研究首要的支撑技术。

（一）双向凝胶电泳的特点

与其他的蛋白质分离技术相比，双向凝胶电泳有以下优势。

（1）一步分离　由于蛋白质组研究要求分离的是细胞所有的蛋白质，因此分离的步骤越少，对样品的影响就越少，而且操作也相对简单。更重要的是，采用其他的分离方法如 HPLC 等，往往要经过多个步骤，每一步骤都会造成样品的损失，特别是对于表达量很少

的蛋白质更是如此，只能适应传统蛋白质研究对一个或一类蛋白质进行分离的要求。蛋白质组是在细胞整体水平上对蛋白质进行研究，要求尽量减少分离过程中样品的损失，只有双向凝胶电泳一次性地将细胞蛋白质分离和展示出来，才能从整体上研究细胞蛋白质的组成和变化。

（2）高灵敏度和高分辨率 凝胶电泳的灵敏度是比较高的，在不同的染色方法下，有不同的灵敏度。通常在银染条件下，灵敏度可以达到 $10^{-15}\sim10^{-18}$ mol 水平，即使是表达量很少的蛋白质，仍然可以检测得到，有利于细胞蛋白质的研究。另外，双向凝胶电泳的分辨率相当高，因为通过以电荷和分子量差异为基础的两次分离，即使是差别很小的蛋白质之间也能得到较好的分离，如可以分辨出糖蛋白的不同糖型成分。良好的分辨率不仅可以保证有效的分离；也是进一步进行分析鉴定的基础，因为分辨不好造成的样品混合，将给分析鉴定乃至数据库检索带来很大的麻烦。

（3）便于计算机分析处理 双向凝胶电泳分离后的图谱经扫描输入计算机，可以进行蛋白质的定位，等电点和分子量的鉴定，以及蛋白质的定量，不同的图谱间可以进行比较。图谱上的蛋白质点经过分析定性后，可以建立蛋白质组数据库。相同细胞的双向凝胶电泳应该有一定程度的重复性，这样只需要比较实验结果和数据库，就可以对电泳分离到的已知蛋白质定性和定量，而不必再重复所有的分析鉴定。

（4）与分析鉴定方法相匹配 双向凝胶电泳可以很好地与分析鉴定方法匹配，电泳后样品可以方便地转移到 PVDF 膜上，进行 N 端或 C 端序列分析。更为重要的是，直接进行蛋白质胶内酶解或膜上酶解目前已经是较成熟的方法，可以很好地与质谱分析匹配。这一点很重要，因为如果不方便做进一步分析鉴定的话，再好的分离也无济于事。

（二）双向凝胶电泳的基本步骤

1. 样品的全息制备

蛋白质样品制备是蛋白质组研究的第一步，无论后续采用怎样的分离或鉴定手段，样品制备都是关键步骤。从蛋白质组大规模研究的角度而言，要求样品制备尽可能获得所有的蛋白质，但由于蛋白质种类多、丰度不一和物理化学特性多样等，要达到真正的全息制备是有难度的。当然，样品制备的方法还必须与后续的分离或鉴定方法相匹配，如采用双向凝胶电泳，需谨慎使用 SDS 抽提蛋白质，如果以液相色谱作为分离手段，太多的去污剂和两性电解质不是一个好的选择。样品制备大体上包括蛋白质的溶解、变性及还原，去除非蛋白质杂质等。虽然看似简单，但实际上样品制备量的多少、溶解的好坏、还原是否完全都将很大程度影响分离的结果。所以，在样品制备方面各种方法是各具特色。在样品制备过程中，一切影响等电聚焦和凝胶电泳的因素都应该考虑在内。样品制备方法可以从简单的缓冲液溶解，到应用离液剂、还原剂及去污剂来进行复合物提取，以及到更复杂的顺序提取（sequence extracting）和亚细胞分级等。由于实验需求的不同以及样品本身的特殊性，没有一种样品制备方法能够广泛地适用于各种各样的样品。所以，对于每一个有着不同要求的样品，都需要通过大量的实验来摸索最合适的实验条件。有效的可重复的样品制备方法是双向凝胶电泳成功的关键。一般来说，一种理想的样品制备方法应该满足以下几点要求：①在合适盐浓度下，可重复地溶解所有的蛋白质，包括疏水性蛋白质，并使样品中的蛋白质以分离的多肽链形式存在；②避免溶解性低的蛋白质（如膜蛋白）在等电聚焦时由于溶解度降低而沉淀析出；③防止蛋白质的化学修饰，包括蛋白质降解、蛋白酶或尿素热分解后所引起的修饰；④排除核酸、多糖、脂类和其他干扰分子；⑤获得的目标蛋白质应在可探测水平，有时需要

去除高丰度蛋白质和不相关蛋白质。

制备得到的蛋白质样品在电泳前需要对蛋白质的浓度进行测定，一般采用 Bradford 法。其他蛋白质定量方法如 BCA 法、Bireut 法和 Lowry 法等都是建立在将 Cu^{2+} 还原为 Cu^{+} 的颜色变化基础上的，而双向凝胶电泳样品制备的裂解液中含有巯基还原剂，会影响这些定量方法的灵敏度。Bradford 法是根据考马斯亮蓝 G-250 和蛋白质的结合来定量蛋白质的，但这种结合也会受到裂解液中的碱性 Ampholytes 和尿素的影响，所以在该法测定前，通常先用 0.1mol/L 盐酸中和样品。

2. 第一向等电聚焦

目前，第一向大多采用 Pharmacia 公司的 IPGphor 或 Bio-Rad 公司的类似仪器，使用商品化的固化胶条，使等电聚焦基本实现自动化，并有良好的重复性。目前通常采用的胶条厚度 0.5mm，宽 3mm，长度有 24cm、18cm、13cm、11cm、7cm 等多种规格，pH 范围有宽窄两类，宽 pH 范围胶条通常为 3～10，窄 pH 范围胶条为 4～7 或 6～11。商品化的胶条是以干胶条的形式和塑料支撑薄膜粘在一起，使用时需在重泡涨液中泡涨。样品可在重泡涨时加入，也可以重泡涨完成后再加，上样量根据胶条的规格及后续采用的蛋白质检测技术来确定，可达毫克级。等电聚焦的电压上升分为三个阶段，首先加上一个比较低的电压，去除胶条中过多的盐离子。然后开始加高压，电压上升的模式为缓慢上升，最后是持续高压，电压上升的模式为快速上升。通常最高电压可达 8000～10000V。总电压时间根据胶条的长度、pH 范围和上样量的多少在一定范围内调整。

第一向等电聚焦结束后，胶条必须经过平衡才能进行第二向电泳。平衡缓冲液的组分一般为 6mol/L 尿素，0.375mol/L Tris pH8.8，2% SDS，30%甘油，2%二硫苏糖醇（DTT，第一次平衡加入）或 2.5%碘乙酰胺（IAA，第二次平衡加入），0.002%溴酚蓝。尿素和甘油可以增加溶液黏度，减少电内渗。SDS 用于变性蛋白质，使蛋白质带上负电荷，这对第二向电泳是十分重要的。DTT 和 IAA 分别用于打开和封闭蛋白质的二硫键。溴酚蓝用来检测电泳的进程。两次平衡时间通常各 15min。

3. 第二向 SDS-PAGE

第二向目前大多仍采用手动装置，与传统的 SDS-PAGE 相比，并没有很大的不同。为了使蛋白质得到更好的分离，使用的胶也越来越大（25cm×25cm），也有同时进行 20 块胶的分离，以积累更多的样品量供分析。

IPG 胶条平衡好后，一般将胶条在电泳缓冲液里浸润一下，这样做一方面可以清洗胶条，去除胶条上的平衡液，另一方面润湿的胶条也容易沿着玻璃板推到 SDS-PAGE 凝胶上端。IPG 胶条转移到 SDS-PAGE 胶有两种方式：一种方式是先将 IPG 胶条放到 SDS-PAGE 胶上，再加入熔化的琼脂糖凝胶液；另一种方式是先加入熔化的琼脂糖凝胶液，再将 IPG 胶条放到 SDS-PAGE 胶上。两种方式各有优劣。第一种方式图谱不易变形，但在两块胶之间容易有小气泡产生，第二种方式非常好地解决了胶之间的气泡问题，但琼脂糖容易凝固，插入 IPG 胶条时有时会造成图谱变形。相比之下，气泡对图谱的影响较小，故我们推荐第一种方式。SDS-PAGE 胶的浓度往往需根据研究目的和样品性质而定，最常用的是 12%和 12.5%的均一胶。采用均一胶的好处是灌胶方便、重复性好，但分子量较高的蛋白质斑点较为聚集，分离不佳。采用 9%～16%的线性梯度胶，分子量不同的蛋白质点在整块凝胶上分布比较均匀，但灌胶不易。电泳过程中温度一般控制在 10～15℃。常见的双向凝胶电泳结果如图 12-1 所示。

4. 蛋白质的检测

电泳后胶上蛋白质的检测有多种方法，灵敏度和分辨率有一些差异，主要根据上样量、样品进一步分析要求等加以选择。不同染色方法灵敏度不同，造成图谱不同。传统的染色方法有考马斯亮蓝染、银染、铜染等。其中，银染以其灵敏度高、分辨率好而应用最广泛。但银染操作较繁琐，而且常常造成蛋白质末端封闭，不利于蛋白质序列分析。银染有 100 多种不同的方法，分为两大类，碱性银氨染色法和酸性硝酸银染色法。酸性硝酸银染色法和碱性银氨染色法相比，它对酸性蛋白质的染色灵敏度稍高，但对碱性蛋白质的染色灵敏度稍低。酸性硝酸银染色法的原理是一个照相的过程，凝胶在酸性环境下与硝酸银温育，银离子附于蛋白质表面，然后在碱性环境中用福尔马林还原成金属银。碱性银氨染色法的原理是用氨水来形成银氨复合物，银氨复合物与胶内 SDS-蛋白质复合物结合，然后在酸性条件下用福尔马林将银氨还原成金属银。由于氨的不稳定性，酸性硝酸银染色法更适合于定量比较分析。一般认为银染的灵敏度是能染 1～10ng 的蛋白质。银染后的蛋白质一般不能用于后续的分析鉴定，但通过染色方法的改进，银染后的蛋白质也可用来分析鉴定。考马斯亮蓝（CBB）染色法操作简单，但所需染色、漂洗时间较长。采用 CBB R-250 染色的灵敏度在 100ng 左右，CBB G-250 的灵敏度稍高于 CBB R-250。还有些实验室采用同位素标记或抗体标记法。放射性同位素标记法显示蛋白质的灵敏度非常高，同位素的掺入是以细胞代谢为基础的，不同蛋白质掺入同位素如^{35}S的量不同，因此有时难以反映蛋白质组中各种蛋白质的实际丰度，从胶上切取蛋白质斑点的不便也是影响它被广泛使用的一个因素，但用来监测蛋白质组表达变化时有其潜力。负染中的锌染和铜染是近年来逐渐受到重视的染色方法，染色后凝胶背景呈不透明的乳白色或青绿色，而蛋白质斑点或条带则是透明的。铜染检测灵敏度稍高于 CBB R-250，锌染的灵敏度则与银染相当，而且与 CBB R-250、CBB G-250 染色一样对蛋白质的后续分析有较好的兼容性。由于金属负染蛋白质是可逆的，具有高灵敏度的锌染在蛋白质组研究中会具较大潜力。

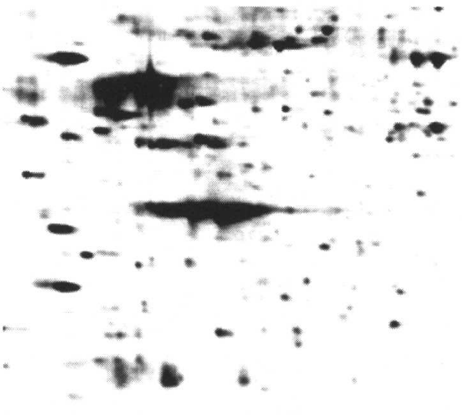


图 12-1　常见的双向凝胶电泳结果照片（化学发光显色）

利用荧光试剂显影凝胶分离蛋白质在蛋白质组研究中的应用越来越广泛。传统的荧光标记往往需要共价修饰蛋白质的自由氨基、羧基或巯基，这种共价修饰的荧光标记法有许多缺点：等电聚焦前标记蛋白质会改变蛋白质的等电点，从而引起蛋白质的不正常迁移；不同蛋白质中荧光基团可修饰的功能基团不同，因此不同蛋白质对标记的灵敏度不同；荧光基团的共价修饰会干扰蛋白质的分离后分析，如蛋白质分子量的轻微改变影响质谱识别鉴定蛋白质。然而，非共价修饰的荧光染料如 sypro red 和 sypro orange 等则与前者不同，它有许多优点。sypro red 和 sypro orange 对聚丙烯酰胺凝胶中蛋白质染色非常灵敏，能检测 1～10ng 的蛋白质，可与银染媲美，且染色的动态范围比银染大，达 3 个数量级；染色步骤简单，一步染色，无需脱色；染色后的蛋白质适于 Edman 降解测序和质谱分析。但图像采集时需 UV 光源或激光扫描仪。一种新型的荧光染色试剂 sypro ruby，其染色原理是利用重金属铷的螯合物来和胶上蛋白质结合，染色专一性好，灵敏度高，而且不会担心染色过度。Lopez 比较了 sypro ruby 荧光染色和银染两种方法，前者的定量重复性（相关系数＝0.9）

要比后者（相关系数=0.8）稍好，银染的线性动态范围为8～60ng，而sypro ruby染色的线性动态范围为1～2ng到1～2μg，sypro ruby的线性动态范围要比银染高700倍左右，而且sypro ruby染色在蛋白质酶解的序列覆盖率和质谱图信噪比方面都优于银染。观察功能相关的蛋白质表达的平行性变化是差异蛋白质组（differential proteome）研究的主要内容，实验需要多次重复，大量双向凝胶图像需要用计算机软件一张一张地分析、比较，使用改进的染色方法或图谱成像方式不仅能简化图像分析，并且可以减少实验工作量。现在已建立了一种称为差异凝胶电泳（differential gel electrophoresis，DIGE）的方法，实际上是采用两种荧光试剂分别标记两个样品，然后混在一起进行双向凝胶电泳。他们用两种花青染料（cyanine dyes）丙基-Cy3和甲基-Cy5分别标记经IPTG（isopropyl thiogalacto side）和未经IPTG诱导的大肠杆菌蛋白质抽提物，两样品混合后在一块胶上双向凝胶电泳，因两种染料分别具有不同的激发光和发射光波长，从一块凝胶上可以采集到代表两种样品的两张图像，经计算机处理，得到一张Cy3/Cy5的比例图（ratio image）（相当于差减图），可以很清楚地看到GAL4VPl6蛋白被IPTG诱导。由Bernhardt等建立的双通道成像（dual channel imaging）是另一种在一张电泳图谱上显示蛋白质组表达变化的方法。在细菌用L-[^{35}S]甲硫氨酸脉冲标记5min后裂解，提取蛋白质，双向凝胶电泳后凝胶经过银染，采集银染的双向图谱，同一块胶再用PhosphorImager采集同位素标记蛋白质的双向图谱。银染图谱经软件转换为绿色的图谱（即保留绿色通道），同位素图谱转换为红色图谱，然后将两图谱合成一张图谱，可以看到新合成的蛋白质为红色斑点，细胞内早已合成的蛋白质为绿色斑点，一直在合成的为黄色斑点。这种方法免去了凝胶配比的麻烦，但操作过程中应戴干净的手套以避免同位素污染。

5. 图谱数字化分析

双向凝胶电泳根据等电点和分子量的不同可一次分离数千种蛋白质，经染色后蛋白质在聚丙烯酰胺凝胶上形成密度不同、分布不均的复杂点图谱。如何对图谱上的这些蛋白质点进行检测、定量、比较、分析和归类是双向凝胶电泳分析软件所要解决的问题。双向凝胶电泳图谱分析一般包括以下几个步骤：图像采集、图像加工、点的检测和定量、凝胶匹配、数据分析、数据解释和构建数据库。图像采集是通过扫描仪或数码相机将凝胶转换为数字图像，其参数设置对于以后的蛋白质点的内在信息采集是比较重要的，因为以后的整个图像加工过程实质上是给图谱上的每个蛋白质点进行精确的定位和定量（以面积和密度作为指标）。图像加工主要是指软件通过一定的算法来去除图谱上非蛋白质点的杂质和染色不同造成的不同深浅的背景。点的检测主要是产生每个点的X轴、Y轴坐标，外形参数和密度参数，并对每个蛋白质点做数字模拟合成。对于点的检测和定量，目前采用的算法可分为高斯拟合（Gaussian fitting）和拉普拉斯算子高斯拟合（Laplacianof Gaussian）两大类。除了在单块胶上获得点的信息外，在一系列不同胶上监测某个蛋白质点的变化是生物学分析所感兴趣的问题。不同图像间的比较首先要确定足够的坐标（landmark），即每块胶上都存在的点。这些坐标用于确定图像之间几何转换的多项式，然后用于其他点的匹配。表达量变化的蛋白质点或已经鉴定的点可在软件中进行数据归类和解释，还可以建立数据库通过因特网来进行成果共享。

双向图谱分析软件的开发始于20世纪70年代末，至今经历了三代的发展历程。最初的程序是使用DEC PDP1Ⅰ家族的小型计算机在VMS操作系统上运行，如Elsie、LIPS、Gellab和Tycho；第二代是基于Unix的程序，盛行于20世纪80年代后期，如Gellab-Ⅱ、

Elsie-4、Kepler、Melanie 和 Quest；第三代是基于 Unix、NT 和 Mac 多种计算机平台的程序，如 2DBioImage、Melanie·Ⅱ和 Phoretix 2D。当今，双向图谱分析软件正向 PC 化发展，而且界面更友好，使用更方便，功能更强大，以适应这一工具进入普通实验室，如 PDQuest、Melanie·Ⅲ、Phoretix 2D 和 ImageMaster 2D 等。近年来随着蛋白质组学的兴起，新的双向凝胶电泳分析软件在自动化方面又有较大改进，如 Z3、Delta 2D 和 Progenesis 等。

6. 操作实例——双向凝胶电泳分离、比较培养肝细胞系的蛋白质组学研究

人肝癌细胞系 HepG2 及人正常肝细胞系 L-02 培养在含 10%小牛血清的 1640 培养液中，生长至 80%左右密度，胰酶消化，离心收集的细胞经预冷的 PBS 漂洗一次，5.48% 山梨醇漂洗三次。去盐后，离心收集于 1.5ml EP 管中，细胞重悬于 150μl 裂解液（8mol/L 尿素，2mol/L 硫脲，4% CHAPS，40mmol/L Tris 和 65mmol/L DTT），超声细胞破碎仪（Soniprep150，美国 MSE）冰浴超声 2min，4℃ 14000r/min 离心 60min，取上清，QUANT 2D（Amersham Pharmacia）试剂盒测定蛋白质浓度，样品置－80℃保存备用。

根据蛋白质定量结果取 1000μg 总蛋白与重泡涨液（8mol/L 尿素，2%CHAPS，0.5% IPG 缓冲液 3～10L，18mmol/L DTT 和痕量溴酚蓝）混合，总体积为 350μl，并吸入 IPG 胶条槽（IPG strip holder）中，IPG 干胶条 3～10L（180mm×3mm×0.5mm）胶面朝下放入胶条槽，并覆盖一层矿物油以防样品液蒸发，盖好盖子后置于 IPGphor Ⅱ 水平电泳仪（Amersham Pharmacia）电极板上，重泡涨和等电聚焦在 20℃ 下自动进行，设置如下：30V，12h；200V，15min；500V，30min；1000V，1h；4000V，1h；8000V，10h。总电压时间积达到 60kV·h 后结束等电聚焦。等电聚焦结束后，IPG 胶条在含 50mmol/L Tris-HCl(pH6.8)，6mol/L 尿素，30%甘油，2%SDS，1%DTT 和痕量溴酚蓝的平衡液中平衡 15min，然后再用 2.5%碘代乙酰胺替换 1%DTT 的同样平衡液中平衡 15min，胶条移至 12%SDS-PAGE 胶上端，0.5%琼脂糖封胶。SDS-PAGE 在 DALT-SIX 电泳槽中进行。10mA/胶电泳 30min 后，换用 25mA/胶，直到溴酚蓝到达胶的底线，剥胶进行染色。

用一片 R350（Amersham Pharmacia）加入 80ml ddH_2O 中，搅拌 5～10min 后加入 120ml 甲醇，搅拌至溶解并过滤，终浓度为 0.2%，此为 R350 储液。将 R350 储液和 20% 冰醋酸等体积混匀，待分析的凝胶放入染色至少 9h 或过夜后，10%冰醋酸脱色至背景干净。结果如图 12-2 所示。

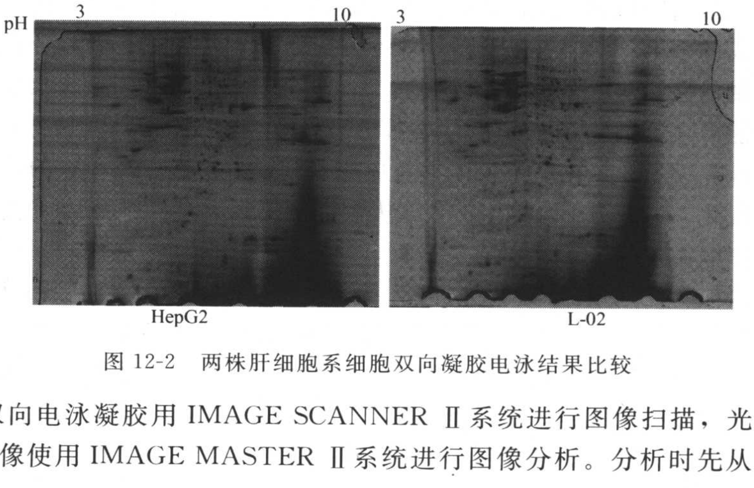


图 12-2 两株肝细胞系细胞双向凝胶电泳结果比较

染色后的双向电泳凝胶用 IMAGE SCANNER Ⅱ 系统进行图像扫描，光学分辨率采用 300dpi。数字图像使用 IMAGE MASTER Ⅱ 系统进行图像分析。分析时先从所有的凝胶图

像中生成虚拟的合成胶（synthetic gel），并以其作为参考胶（reference gel），建立匹配并分析。差异的蛋白质点可通过挖胶回收，按质谱分析程序进一步鉴定蛋白质的种类和序列。

（三）蛋白质组分离的非凝胶技术

在蛋白质组学的研究技术中，双向凝胶电泳（2D-PAGE）由于具有无可比拟的高分辨率而始终占据着重要的地位。然而，这种技术本身也存在一些难以克服的缺陷，如在双向凝胶电泳图谱上，并非每个蛋白质点所包含的蛋白质量都足以用于鉴定蛋白质；操作费时费力，难以实现与作为后续分析鉴定手段的质谱直接联用，不易自动化；存在偏向性问题，如对极大蛋白质（分子量＞200000）、极小蛋白质（分子量＜8000）、低丰度蛋白质（如信号蛋白和转录因子）、极碱性蛋白质和疏水性蛋白质（膜结合蛋白质和跨膜蛋白质）难以进行分离分析。这些缺陷都在一定程度限制了双向凝胶电泳的应用。因此，发展新型高分辨率、高通量、高峰容量的分离分析方法已经成为加速蛋白质组学研究的关键。近年来，分离科学发展出了一些新型的分离模式，使高效液相色谱、毛细管电泳和毛细管电色谱等技术在蛋白质组学研究中焕发出了新的活力。这些新技术，有的是不以凝胶为分离介质的，我们暂且将其归为蛋白质组学研究中的非凝胶技术。这些非凝胶分离技术所表现出的高效、快速、高自动化程度、高灵敏度、低检测限等优点已经极大并有可能进一步推动蛋白质组学的相关研究。

1. 非凝胶色谱技术

（1）一维色谱与质谱联用技术　高效液相色谱经过几十年的发展，已经派生出多种分离模式。但由于在蛋白质组学的后期研究，即质谱的鉴定分析当中最常用的质谱离子源是电喷雾离子源（ESI）和基质辅助激光解吸离子源（MALDⅠ），这两种离子源都要求整个质谱系统中最好不被引入过高浓度的盐，所以在将一维色谱和质谱联用时，一般是将体系中没有盐存在的反相高效液相色谱（RP-HPLC）和质谱系统联用，从而达到最终的分离鉴定目的。比较方便而简单的联用技术是直接在电喷雾质谱里喷雾针所用的毛细管内填充或原位聚合一段反相液相色谱的填料，实现分离和喷雾一步完成。在某种程度上，这种技术相当于将高效液相色谱“整合”到质谱系统中。采用这种技术，多肽的最低检测限达到了 fmol 级。目前，内部装有填料的喷针已经实现了商品化。

（2）多维色谱及其与质谱联用技术　与一维分离模式相比，多维分离技术的最大特点是可以极大地提高峰容量。1995 年，Giddings 等指出峰容量的提高程度和样品组分分布的有序程度有着密切的关系，并引入了样品维数（sample dimensionality）——能够精确识别样品中所有组分的变量数的概念。他们认为，当样品维数大于系统维数时，组分的峰分布是无序的，会限制分离的效果；而当样品维数小于系统维数时，尽管组分的峰分布是有序的，但多维系统高峰容量的性能并没有得到充分的发挥；只有当样品维数和系统维数相等的时候，多维系统的分离能力才能得到最大程度的发挥。同年，Giddings 构建了二维分离模式下实际峰容量的数学计算模型，该模型不仅能够拓展应用到评价多维分离模式的性能，也可以应用于系统的条件优化。另一方面，Murphy 等从理论和实验上对二维分离模式下第二维的采样速度对分离的影响做了深入研究。结果表明，进入第二维分离的采样时间越短，整个系统的选择性越高。目前广泛采用的二维色谱包括离子交换色谱-反相液相色谱、体积排阻色谱-反相液相色谱、反相液相色谱-反相液相色谱、亲和液相色谱-反相液相色谱等。这些二维色谱技术为蛋白组学的研究提供了强有力的支持。而一些研究小组为了更充分地发挥液相色谱技术在蛋白质组学研究中的优势，已经开始尝试三维色谱的联用，如 Moore 等尝试了体积排阻色谱-反相液相色谱-毛细管区带电泳的三维分离技术，MacCoss 等成功实现了反相液相色

谱-离子交换色谱-反相液相色谱的三维联用技术等。

2. 非凝胶电泳技术

主要包括液相等电聚焦技术及毛细管电泳技术。

(1) 液相等电聚焦技术 采用无载体系统，仅在液相中进行等电聚焦，不存在从载体上回收样品等问题。普通的 2-D 电泳第一向等电聚焦是在固化的胶条上进行，而在液相中进行等电聚焦有许多优点，首先该技术能够大量纯化蛋白质和多肽，尤其在相当高的上样量时也具有相当高的分辨率；其次可在非凝胶介质上根据等电点将蛋白质的复杂混合物分级，每一级都能被收集，并且通过电泳或色谱进一步分离或分析；生物样品的活性易于保持；对于低丰度蛋白质和低分子量的蛋白质也具有很好的分离效果；可以避免在 2-D 凝胶电泳中存在的重复性差的问题。目前，能够将复杂混合物分级的液相 IEF 仪器有多种：①Rotofor cell 仪器（RotoforCell Apparatus)；②多腔等电膜电解器（multicompartment electrolyzer with isoelectric membranes)；③Bier 首创的循环等电聚焦仪（recycling isoelectric focusing，RIEF；RF3)；④自由流等电聚焦（free-flow IEF，FFIEF)。在 RIEF 中，样品介质通过冷却室循环，而 FFIEF 中样品被持续注入作为载体的两性电解质溶液中。复杂的蛋白质溶液很容易被分级为很多组分，它们再被收集，并应用正交技术进一步分级。除了上述 RF3 以外，还有其他一些用于循环等电聚焦的仪器来分离蛋白质混合物，如 Tananahalli 等所应用的两种等电聚焦仪，其自动化程度更高，在分离过程中实现了计算机控制，并且各个组分的 pH 和浓度也可以在线检测，使得实验过程更加精确。使用 RF3 分级后，再把蛋白质溶液在反相液相色谱中进行分离，也是一个很好的策略，经液相等电聚焦和反相色谱分离后，使用一种软件可以把蛋白质模拟成二维胶的分离模式。

(2) 毛细管电泳技术 与等电聚焦类似，毛细管电泳也是一项具有很高分离效率的分离技术。20 世纪 90 年代，这种技术被广泛应用于 DNA 分析。人类基因组测序最终能够提前完成，很大程度上也要归功于这种分离技术。毛细管电泳的基本原理是在直径 25～100μm 的弹性石英毛细管中进行电泳，毛细管中填充了缓冲液或凝胶。与平板凝胶电泳相比，毛细管电泳可以减少热的产生，缩短分析时间并提高分辨率，电泳所需样品体积很小，且多种分离模式能改变选择性，扩大应用范围。目前，在蛋白质组学研究中应用较多的是毛细管区带电泳、毛细管等电聚焦等。多维毛细管电泳、毛细管电泳芯片等在蛋白质组学中的应用正处于起步阶段。

3. 毛细管电色谱技术

毛细管电色谱是一项新兴的微分离技术，最初是通过在毛细管内填充液相色谱填料，在毛细管两端加以高电压，以电渗流驱动流动相，根据溶质在固定相上的保留强弱和电泳淌度的大小不同而实现分离。由于毛细管内有色谱固定相，电渗流驱动的流动相流型是楔子型的，因此这种微分离技术在具有高效液相色谱的高选择性的同时，又具有毛细管电泳的高效性，已成为分析化学界近年来比较关注的领域。根据毛细管内固定相种类的不同，毛细管电色谱又有开管、填充和原位柱之分。毛细管电色谱诞生之初的应用领域主要是对中性小分子化合物的分离，但是随着这种技术在方法学上的逐步完善，更多的分离模式和固定相被相继采用。当前，毛细管电色谱在生物学领域内也有了初步的应用，但只局限于比较简单的样品。蛋白质和多肽等生物样品，表面带电情况比较复杂，而毛细管电色谱的固定相表面又是带有一定的正电荷或者负电荷的（否则电渗流将无法产生)，两者之间可能会存在一定的吸附作用。因此，尽量减少甚至避免这种静电相互作用就成为毛细管电色谱分离生物样品的关

键。从目前对电色谱在多肽和蛋白质领域内应用情况来看，虽有一些成功分离小肽和标准蛋白质的报道，但该技术还需进一步发展，才能适应蛋白质组学研究的需要。

二、蛋白质组分析技术

相对分离技术而言，蛋白质组分析技术有较多种的选择，比如质谱方法、蛋白质序列分析、氨基酸组成分析等。其中质谱分析和蛋白质序列分析最为重要，而质谱以其快速、准确、灵敏而成为蛋白质组主要的分析技术。

蛋白质组中的分析技术与常规蛋白质研究的分析技术也有所不同，常规蛋白质研究通常是解决未知序列或验证已知序列，蛋白质序列分析是最可靠的方法。但蛋白质组研究要求在短时间内分析双向电泳分离到的上千种蛋白质，即使是丰度很低的蛋白质也希望被鉴定。因此，分析方法的速度、灵敏度和准确性至关重要。传统的序列分析测定一个蛋白质的 N 端 10 个氨基酸残基需要至少 5h，而且若蛋白质末端被封闭，将无法测序。而质谱技术可 1h 内完成一个蛋白质整个肽谱的鉴定，得到完整的蛋白质全序列，经计算机数据库查询，可以很快地鉴定蛋白质。另外，对于末端封闭或某些残基被修饰的情况，质谱方法几乎是目前惟一最快速有效的方法。因此，质谱技术理所当然成为了蛋白质组分析方法的首选。质谱技术在蛋白质组研究中的作用有以下几方面。

1. 肽质量指纹图谱和肽序列分析

蛋白质经过双向电泳后，分离到的蛋白质点被切割下来，进行胶内的酶解，或转移到 PVDF 膜上，进行膜上酶解。这两种方法各有千秋，目前以胶内酶解应用更多。酶解后的产物可用 LC/MS 分析肽质量指纹图谱，进而采用 MS/MS 测定所选肽段的序列。测得的肽质量指纹图谱与数据库中已知蛋白质的肽质量指纹图谱库进行比较，就可以确定蛋白质是否已知，或是何种蛋白质。这个过程比常规的 Edman 降解快得多，因此很适合大规模的蛋白质组研究，这也是质谱技术在蛋白质组研究中占据重要地位的一个原因。要强调的是，在这方面计算机数据库起了很大的作用。对于 MALDI-TOF-MS，可以采用膜上酶解，然后直接对 PVDF 膜上的样品做肽质量指纹图谱分析，MALDI-TOF-MS 在使用了双向凝胶电泳(2-DE)技术以后，使得肽质量指纹图谱精确度大大提高，已经可以仅仅根据质量指纹图谱查询数据库，获得蛋白质信息，非常快速简单。而另一方面，由于电泳中会引入各种小分子物质如 SDS、CHAPS 等，对质谱分析有一定的干扰，特别是会影响电喷雾质谱的喷雾效果。虽然 MALDI-TOF-MS 对小分子杂质的耐受性较好，但无论如何，电泳后酶解前小分子杂质的去除仍然十分重要，是影响分析效果的关键因素之一。不过，在肽序列分析上，质谱法测序对于未知序列肽段的结果还不肯定。质谱法目前还不能够完全取代 Edman 降解法，而两者互补往往可以得到确定的结果。

2. 具有翻译后修饰的蛋白质鉴定

翻译后修饰包括磷酸化、糖基化、N 端封闭等。蛋白质一级结构可以从基因序列演绎，但翻译后修饰的信息几乎无法仅从基因序列得到。而众所周知的是，这些翻译后修饰特别是磷酸化、糖基化对于蛋白质的功能乃至细胞的功能都十分重要，因此这些修饰的鉴定对于蛋白质组特别是功能蛋白质组的研究意义重大。Edman 降解法对于这些修饰的鉴定无能为力，而采用质谱法可通过特征离子监测的方法很快确定磷酸化肽，通过串联质谱确定磷酸化位点。在糖蛋白分析方面，质谱更显示了独特的优势，不仅可以通过质谱、蛋白酶解和糖苷酶酶解相结合的方法寻找糖肽，鉴定糖基化位点，还可依靠 MS/MS 和 MS_n 分析糖链组成、

结构甚至分支情况。翻译后修饰的快速鉴定大大简化和加速了蛋白质组的研究进程。

3. 质谱技术在蛋白质组研究中的其他作用

质谱技术在蛋白质组研究中的主要作用是鉴定蛋白质，此外还有一些其他用途，如蛋白质二硫键的定量和定位、蛋白质-蛋白质相互作用的分析、蛋白质与其他分子得到相互作用以及蛋白质的二级结构的分析等。这些分析将在蛋白质组研究进行到一定阶段，获得了有价值的蛋白质（如正常细胞与异常细胞的差别）后，再对单个蛋白质进行研究时发挥作用。

三、蛋白质组数据库的建立和生物信息学

蛋白质组研究最重要的一个方面就是数据库的建立，数据库的范围很广，包括蛋白质序列数据库、质谱数据库、双向电泳图谱数据库等，几乎所有有关蛋白质结构的数据库都被应用到蛋白质组的研究。蛋白质经分离分析后必须通过数据库的查询，才能确定蛋白质的性质，是已知或未知、有否翻译后修饰、与同类蛋白质同源性如何等。互联网上目前有很多这种数据库可供查询。而狭义上的蛋白质组数据库指的是双向电泳图谱数据库。对于某种细胞的蛋白质组而言，双向电泳图谱数据库的建立是指蛋白质被分离和分析后对图谱中的每一个蛋白质点进行定位和定性，经计算机处理后得到的二次图谱。一个完整的蛋白质组研究包括细胞蛋白质的抽提，双向电泳的分离，分离图谱的计算机扫描、定位、定量和等电点的确定，蛋白质酶解并从胶上抽提出来，质谱结合其他分析方法对蛋白质的分析，分析后对蛋白质的鉴定，最后得到计算机二次图谱，也就是双向电泳数据库的建立。因此，每一种细胞数据库的建立和不断完善，意味着越来越多的蛋白质得到分离和分析，是蛋白质组研究取得进展的标志。数据库建立后可以进行实验室内的比较，比如正常细胞和癌细胞之间的比较，这种比较在疾病研究和药物开发中很有意义。数据库还能通过互联网进行实验室之间的比较，但这通常较为困难，因为不同实验室方法上哪怕是微小的差异，也会使同一样品的分离图谱有区别。在各种蛋白质组研究网站上，一部分细胞已经建立了数据库。

目前最大也最著名的蛋白质组数据库是 SWISS-2DPAGE，始建于 1993 年，到 1999 年 9 月为止，经过了 9 次更新，建立了 23 种人和鼠细胞的蛋白质组数据库，另外还包括大肠杆菌、酵母等。专门的大肠杆菌数据库 ECD2DBASE 目前已更新到第 6 版，分离到 1600 多种蛋白质。

可以看到，蛋白质组研究的整个过程实际上都与生物信息学密切相关，无论是双向电泳图谱的分析，还是质谱数据的解析，尤其是最终蛋白质组数据库的建立，实验室间的相互比较，都依赖于生物信息学的建立和发展。

四、蛋白质组研究的发展趋势及应用前景

蛋白质组学一经出现，就有两种研究策略。第一种策略可称为“竭泽法”，即采用高通量的蛋白质组研究技术分析生物体内尽可能多乃至接近所有的蛋白质，这种观点从大规模、系统性的角度来看待蛋白质组学，也更符合蛋白质组学的本质。但是，由于蛋白质表达随空间和时间不断变化，要分析生物体内所有的蛋白质是一个难以实现的目标。第二种策略可称为“功能法”，即研究不同时期细胞蛋白质组成的变化，如蛋白质在不同环境下的差异表达，以发现有差异的蛋白质种类为主要目标。这种观点更倾向于把蛋白质组学作为研究生命现象的手段和方法。

早期蛋白质组学的研究范围主要是指蛋白质的表达模式（expression profile），随着学

科的发展，蛋白质组学的研究范围也在不断完善和扩充。蛋白质翻译后修饰研究已成为蛋白质组研究中的重要部分。蛋白质-蛋白质相互作用的研究也已被纳入蛋白质组学的研究范畴。而蛋白质高级结构的解析即传统的结构生物学，虽也有人试图将其纳入蛋白质组学研究范围，但目前仍独树一帜。

蛋白质组学的研究技术目前发展较快，主要是质谱鉴定技术的发展和新的分离手段的应用，如二维色谱、毛细管电泳技术等。蛋白质组学的研究方法将出现多种技术并存、各有优势和局限的局面，而难以像基因组研究一样形成比较一致的方法。除了发展新方法外，更强调各种方法间的整合和互补，以适应不同蛋白质的不同特征。另外，蛋白质组学与其他学科的交叉也将日益显著和重要，这种交叉是新技术、新方法的活水之源，更重要的是，蛋白质组学与其他大规模科学如基因组学、生物信息学等领域的交叉，所呈现出的系统生物学（systems biology）研究模式，将成为未来生命科学最令人激动的新前沿。

由于蛋白质与细胞的生理功能直接相关，蛋白质组的研究应用前景非常广阔。在基础研究方面，近两年来蛋白质组研究技术已被应用到各种生命科学领域，如细胞生物学、神经生物学等；在研究对象上，覆盖了原核微生物、真核微生物、植物和动物等范围，涉及到各种重要的生物学现象，如信号转导、细胞分化、蛋白质折叠等。已开展的研究包括研究细胞不同生长时期的蛋白质组、正常细胞与异常细胞蛋白质组的差别等；在应用研究方面，蛋白质组学将成为寻找疾病分子标记和药物靶标最有效的方法之一。在对癌症、早老性痴呆等人类重大疾病的临床诊断和治疗方面蛋白质组技术也有十分诱人的前景。

第三节　蛋白质组学技术平台与完整解决方案

蛋白质组学的发展依赖于双向电泳技术的发展、电喷雾质谱（ESI-MS）和基质辅助激光解吸飞行时间质谱（MALDI-TOF-MS）技术的发明以及在蛋白质分析中的应用和蛋白组信息学的兴起。蛋白组学技术平台是将先进的蛋白组学研究系统与生物信息学平台、基因组学技术平台有机地结合在一起，不仅在研究水平上而且在规模上进行蛋白组学的研究，成为蛋白组学研究重要发展方向。

一、蛋白质组学技术平台

蛋白组学平台采用全方位、高通量的技术路线，包括双向电泳技术、MALDI-TOF 质谱技术、LS/MS/MS 和蛋白质芯片等进行蛋白质组学的研究，并结合 ICAT 等灵敏的生物检测技术，形成规模化的重大疾病、重要生物资源的蛋白质组学研究技术平台。在此基础上，可建立独特的蛋白质组数据库，研究疾病的相关蛋白和具有重要应用前景的生物标记分子，建立疾病的早期诊断和治疗监测方法。从蛋白质与基因组双重水平上研究疾病和重要生物资源。

该平台包括蛋白质分离纯化技术装备、免疫学技术装备等（如图 12-3 所示），如先进的层析系统、电泳系统、酶标仪、紫外可见光分光光度计等设备，并结合高通量蛋白质表达技术平台，如微阵列、生物芯片等大规模筛选技术，大量、快速地实现 cDNA 库、多基因或

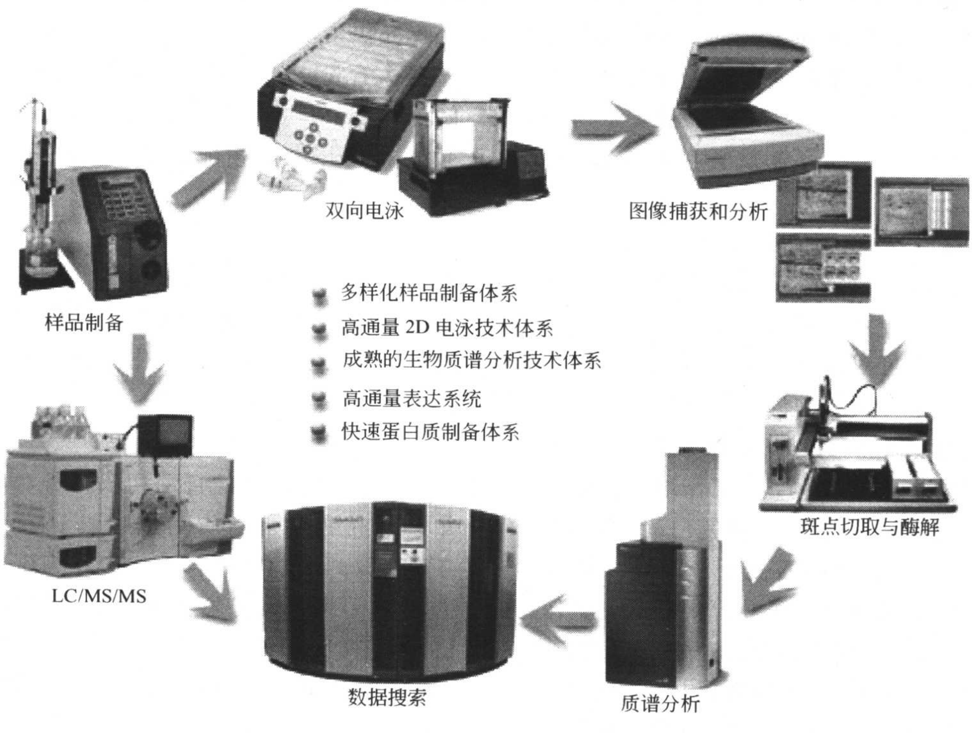


图 12-3　蛋白组学技术平台主要设备及流程

片段的表达，筛选各种抗原、受体、细胞因子等重组多肽或蛋白质。

二、蛋白质组学研究的完整解决方案

以蛋白质为研究重点的蛋白质组学研究已经成为当今生命科学的重点研究方向之一。然而，由于蛋白质的分离和鉴定非常费时，目前测定蛋白质的技术远远落后于破译基因组的进展，最好的实验室每天只能分离和识别出 100 种蛋白质。据估计，人体内可能有几十万种蛋白质，这大概需要 10 年时间进行识别。

为了加快蛋白质组学研究进程，许多世界著名的生物技术公司倾力发展蛋白质组学研究设备，已经开发了完整的蛋白质组学解决方案。这些系统包括由一系列机械手臂与软件，并结合了二维电泳实验设备与质谱仪，可以进行高效、自动化且具重复性的实验分析。在这些技术平台上，研究工作更富成效，重复性更好，各仪器之间配合无隙，其整合性及标准性使得研究进程大大加快，原来需要 9～12 个月才能获得数据结果减少到只需要 9～12 周。一套完整的系统具备蛋白质组研究所需的众多功能：2D 电泳、图像获取、2D 胶分析、蛋白样品切割、蛋白消化、MALDI 样品准备、消化及点样、数据分析整合，再加上制备好的胶、试剂及附件，使研究工作可以立即展开，成为进行蛋白质组学研究的利器，大大加速了蛋白质分离和鉴定的速度。该系统主要由以下几部分组成。

1. 2D 电泳系统

该系统主要进行 2D-PAGE 第一向等电聚焦凝胶电泳和第二向 SDS-PAGE 电泳，设备包括 2-D 电泳系统所需的各种设备，如 IPG 胶条电泳、管状制胶设备、二维电泳装置、电源设备、半导体冷却器及各种相关的蛋白纯化试剂盒。

完整的2D-PAGE电泳所需的各种设备，使电泳更加简便，大大节约研究时间。第一向等电聚焦凝胶电泳和第二向SDS-PAGE大面积板胶提供清晰的电泳图像，有效提高单体、磷酸化和糖基化蛋白的分离，使其分辨率更高。具备大容量，如可同时容纳15块1mm一维管状胶，或8块2～3mm管状胶，10块IPG胶条和10块二维电泳板胶。适用各种胶，如管状胶、IPG胶条、预制胶、自制胶和SDS PAGE胶等，并提供相配套的超纯化学试剂和药品。

2. 蛋白凝胶成像系统

蛋白凝胶成像系统提供高灵敏度、高分辨率的大面积蛋白凝胶成像和分析。由于采用高灵敏度、高分辨率数字成像系统及冷CCD数码相机，大大提高图像信噪比；并提供白光和紫外光源，可对银染、荧光、胶片、考马斯亮蓝、EB染色图像进行数字化分析，配合各种蛋白质组学分析软件进行各种图像分析。

3. 蛋白质点自动切取机器人平台

自动切取机器人平台工作站主要用于2D蛋白凝胶上蛋白质点的自动切取，可对考马斯亮蓝、银染、荧光染色的蛋白凝胶进行操作，并进行高通量蛋白质筛选。如Investigator ProPic工作站可将2D凝胶成像分析和蛋白质点自动切取两个功能整合在一起，为蛋白质组学研究的后续分析采集和提供样品。工作站通常包括密闭暗箱及机器人手臂，切割精度高，内置可换光源（白光和紫外光源），能容纳8块96微孔样品板的样品盘，并装配Gilson泵系统和高分辨率冷CCD数码相机的成像系统。凝胶成像和蛋白质点切取功能的整合，使得成像后凝胶不必移动而原位切取，避免移动过程中产生的位移和污染；切取结束后，可再次原位成像，检查切取结果准确性；可对操作过程实时监控。切取速度达120个/h；全封闭的操作环境，有效防止角蛋白污染；超声水浴清洗和泵冲洗两种方式清洗枪头，有效防止样品交叉污染；污染最小化；内置水合装置，有效避免操作过程中因凝胶变干而影响切取准确性和样品回收率。

工作站配置蛋白凝胶图像分析软件、自动控制专用软件（控制数码相机曝光和切割机器人的采样操作），可在数小时的操作周期内无需专人管理，完全自动化，结果可靠。

4. 蛋白凝胶图像分析专业软件

该软件适用于大规模、高通量蛋白质组学研究，能快速处理大量2D蛋白电泳凝胶图像，能够对蛋白质点进行自动识别、定量和比较；自动计算2D凝胶上蛋白质点的分子量和p*I*；自动生成蛋白表达量变化曲线或柱形图；能进行半自动或全自动蛋白质点匹配，不需另外提供标记；可对多块相应凝胶上的蛋白质点数据进行T-检验或其他统计分析；数据可以传送给自动工作站进行蛋白质点自动切割；有些还带有本地计算机数据库系统，可以方便地进行蛋白研究的数据查询和数据管理。

5. 自动化蛋白酶解工作站

自动化蛋白酶解工作站是进行全自动蛋白酶解的理想设备，适合于从2D或SDS-PAGE胶上切取胶块的自动化酶解。通过进行反应程序的设定，实现酶解过程中的自动清洗、加酶和恒温孵育步骤循环。由于采用自动化操作和温控反应系统，精确控制酶解反应条件，节约了反应时间，提高反应效率。

6. 自动化MALDI点样工作站

该系统用于蛋白质质谱分析前的样品浓缩、脱盐和MALDI点样，点样的多肽可以进行后续的质谱分析，适用于大多数的MALDI标靶。

7. 自动化蛋白质酶解点样工作站

酶解点样工作站将能够自动完成凝胶蛋白质点酶解及酶解后样品的浓缩、脱盐和 MALDI 点样等一系列过程，可多个样品针头同时操作使得实验结果高效和自动化。该仪器简化了蛋白质组学研究的烦琐步骤，大大加快了研究的实验进程。酶解后的样品适用于多种后续分析，包括 LC（液相色谱）和 CE（毛细管电泳），尤其适用于质谱分析，如 MALDI-TOF 和 MS/MS。该系统的酶解和 MALDI 点样两个操作过程可以分开单独进行，也可以整合成一个完整的过程进行自动化操作。

8. 快速数据管理和检索系统

许多蛋白质组平台系统，如著名的 Genomic Solutions Inc. 公司产品——RADARS/MS 快速数据管理和检索软件系统，是一个完整的蛋白质组学研究智能平台。它整合了客户机/服务器、数据库和外联网（extranet）应用软件，可以方便地进行蛋白质的分析鉴定和传输，提供最快、最可信和最完整的蛋白鉴定解决方案。通过 RADARS/MS 系统可以方便分析大量的蛋白质组学研究数据，快速发现和确认药靶及疾病标志物，适合于蛋白质组学研究、制药开发研究等。

总之，蛋白质组学平台的系列仪器及软件为蛋白质组学研究提供了完整的解决方案，大大加快了研究进程，提高了实验效率。用户可以根据自己现有的设备情况，有选择性地配备相应的仪器及软件，形成适合自己的一整套完整的蛋白组学研究系统。

（屈　伸）

第十三章　生物芯片技术

人类基因组计划（HGP）的实施以及后基因组计划的启动，不仅破译了编码人类基因组的核苷酸序列，还对蕴藏在序列中的信息内涵进行了发掘，同时也推动了包括生物信息学、基因组学、蛋白组学等新型学科的迅速发展，大规模测序、寻找新基因及基因组功能研究相关的生物技术也应运而生。生物芯片技术，特别是基因芯片和蛋白质芯片的问世，被称为是震动生命科学界的一次“迷你”革命，它们将深刻地改变着生命科学的发展乃至人们的生活和工作。本章将就生物芯片的原理、技术及应用做一简单的介绍。

第一节　概　　述

一、生物芯片的概念和特点

生物芯片技术就是近年发展起来的一项高效率、高通量的生物样品检测技术，是大规模获取生物信息的重要手段。生物芯片技术为人类基因组学从理论研究向实用研究过渡以及生命科学从分子水平研究向细胞乃至整体水平研究的回归架起了一座桥梁。

什么是生物芯片？和微缩集成电路一样，生物芯片也是在一块指甲大小的硅片或尼龙膜（凝胶、金属）等材料上，将生物分子探针（包括细胞、蛋白质、核酸及其他生物组分）以大规模阵列的形式排布，形成可与目的分子相互作用、并行反应的固相表面。将生物芯片与荧光标记分子进行化学反应（如杂交等），经过激光扫描，不同反应强度的标记荧光将呈现不同特征的荧光发射光谱，用 CCD 相机（charge couple device）或激光共聚焦显微镜收集信号后，采用计算机软件分析数据和处理结果，从而获得大量的相关生物信息。

生物芯片对样品的检测是以高通量、集成化、并行化和微型化为特征。举例来说，要研究肿瘤细胞的基因表达发生了哪些变化，或想了解在炎症时机体和细胞的反应性变化，就必须对肿瘤细胞的相关基因（如癌基因、抗癌基因、凋亡基因、细胞周期相关基因等）进行逐一分析，或对炎症相关的细胞因子、趋化因子、黏附分子等的表达状况进行研究。如果这些工作采用常规的分析方法来做，将耗费大量的人力和物力去从事重复性的工作。一般而言，常规做法将会采用各种不同的方法在基因水平、mRNA 转录水平或蛋白质翻译水平上进行研究。如果想观察 mRNA 表达水平的改变，首先就要提取细胞的总 RNA，然后通过逆转录反应，获得 cDNA，再以其为模板进行定量 PCR 扩增，以确定不同基因表达水平变化；也可采取 Northern 杂交的方法进行分析；不管怎样，总是一个基因一个基因地、不连续地去分析，效率之差显而易见。如果以一天分析一个基因的速率计算，分析 100 个基因至少也得花去 3 个月左右；加上不是并行地进行操作和分析，实验的组间误差和批间误差难免产生。但是采用生物芯片来研究基因的表达，则可以避免这些问题，因为它具有高通量、集成化分析的优势，特别适应基因组

和后基因组时代大规模生物信息分析的需要。尤其是新近发展的“缩微实验室芯片（lab on a chip)”，通过微细加工工艺在芯片上制作微滤器、微反应器、微泵、微阀门、微电极等结构元件，以实现将不连续的化学分析过程移植到生物芯片中，这样准备出来的芯片就像一个全功能的实验室，包含所有的分析平台，从样品制备、生化反应、功能检测到结果分析，一气呵成，极大地缩短的检测和分析时间，节省了实验材料，更突出体现生物芯片的连续化和自动化的优势和芯片技术发展的方向。

二、生物芯片的分类

和所有的分类方法一样，根据不同的标准同一种芯片就有不同的归类组别和定位。最初的生物芯片，仅用于 DNA 序列的测定、cDNA 表达谱分析、基因突变的检测和鉴定等，故称之 DNA 芯片或基因芯片。目前，生物芯片的内涵已扩展和覆盖至检测免疫反应、受体结合、小分子捕获、药物筛选等众多领域，种类繁多，分类也趋于更详细和全面。下面将各种不同的分类方式作简单介绍。

（1）根据基片上交联固定的识别分子来分类　可分为基因芯片、蛋白质芯片、肽芯片、细胞芯片、组织芯片、寡核苷酸芯片等。

（2）根据芯片的功能来分类　又可分为测序芯片、基因作图芯片、基因表达谱芯片、克隆选择和文库筛选芯片、物种鉴定和基因组分型芯片、突变检测芯片、多态性分析芯片、药物芯片、噬菌体文库筛选芯片等。特别值得一提的是能将样品分离、纯化、提取、混合、反应、检测等功能集合于一体的微缩实验室芯片（lab on a chip)，将代表微流路生物分析技术发展的最高水平。

（3）根据芯片的用途来分类　可分为分析芯片、检测芯片和诊断芯片。

（4）根据生物化学反应过程和阶段来分类　可将生物芯片分为不同的类型，例如用于样品制备的生物芯片，用于进行生化反应的芯片及各种检测用的生物芯片。

样品制备芯片的目的是将通常需要在实验室进行的样品制备的多个操作步骤集成于微芯片上。目前，样品制备芯片主要通过升温、变压脉冲以及化学裂解等方式对细胞进行破碎，通过微滤器、介电电泳等手段实现生物大分子的分离。

生化反应芯片在芯片上完成生物化学反应。例如，在芯片上进行 PCR 反应的 PCR 芯片。由于检测和分析的灵敏度的需要，通常在对微量核酸样品进行检测时必须事先对其进行一定程度的扩增，PCR 芯片的诞生无疑为快速扩增样品及大量扩增多个 DNA 片段提供了有力的工具，还可以节约实验试剂。

检测芯片是用来检测生物样品质或量的一类芯片。例如，用毛细管电泳芯片进行 DNA 突变的检测；用 DNA 微点阵芯片检测 cDNA 表达谱；用寡核苷酸芯片进行突变分析核基因多态性；用蛋白质或多肽微阵列芯片（也称蛋白质或多肽芯片）检测大量不同的蛋白质和进行抗原表位分析。

有的生物技术公司还将所有的生物芯片按照以下方法进行分类（其中并没有将药物芯片单独归类）。

- 生物芯片
  - 信息生物芯片
    - 基因芯片
      - DNA 芯片
      - 寡核苷酸芯片
    - 组织芯片
    - 蛋白质芯片
  - 功能生物芯片
    - 微流体生物芯片
    - 微缩芯片实验室

第二节 基因芯片

一、基本原理

基因芯片在一微小的基片（硅片、玻片、塑料片等）表面集成了大量的 DNA 分子识别探针，能够在同一时间内平行分析大量的基因，进行大信息量的筛选与检测分析。这些 DNA 分子探针可以是基因组 DNA、cDNA 或脱氧寡核苷酸片段等，用于构成不同类型的基因芯片，如用于基因多态位点（SNP）、基因突变检测、基因测序的寡核苷酸芯片；用于定量监测大量基因表达水平的 cDNA 表达谱或 EST 检测芯片；用于基因作图、物种鉴定和进化分析的 DNA 芯片等。

基因芯片的工作原理与经典的核酸分子杂交方法（Southern 杂交、Northern 杂交）是一致的，都是应用已知核酸序列作为探针与样品中未知的互补靶核苷酸序列杂交，通过随后的信号检测进行定性与定量分析。

二、基因芯片的制备和检测

基因芯片的设计实际上是指在一张表面修饰好的基片上，根据不同的实验目的，选择、排布核酸探针的策略（包括战略和战术）和艺术。在此前提下，然后进行下一步的具体操作。基因芯片主要技术流程包括：基片的选择与芯片制备、靶基因的标记、芯片杂交与杂交信号检测、杂交信号的分析和生物信息的获取等，这些内容都会在下面作详细的介绍。

（一）芯片载体的表面修饰

（1）芯片载体　用于连接、吸附或包埋各种生物分子，并使其以固相化的状态进行反应的固相材料统称为载体。载体通常是片状或膜状的。

（2）载体材料　作为载体材料必须符合下列要求：①表面具有多种活性基团，便于与各种生物分子连接；②具有良好的光学特性，以便于透射光或反射光的测量；③单位载体表面上有最佳结合容量；④物理和化学的惰性材料，不干扰表面生物分子的功能和化学反应；⑤足够的机械强度，能耐受一定的压力；⑥具备良好的生物兼容性。

载体材料有 4 类，即无机材料、天然有机聚合物、人工合成的高分子聚合物、高分子聚合薄膜。目前，常用的有玻璃片、金属片，高分子材料薄膜（如尼龙膜）等，但以玻璃片和尼龙膜应用最为广泛。

（3）载体修饰　即载体的活化。通过化学反应在载体表面键合上各种各样的活性基团，形成具有不同生物特异性的亲和表面，以共价偶联不同类型的生物分子，如蛋白质、核酸、多肽等。

（4）修饰载体的选择　芯片载体的修饰有很多方法，如二硫键法、异硫氰酸盐键法等。最常用的两种玻片，分别用醛基或氨基修饰。

醛基修饰的玻片主要利用 Schiff 碱结合的原理。在中性室温条件下，玻片上的醛基与氨基反应生成弱 Schiff 碱而将待加样的分子固定下来。这种玻片可以制备 cDNA 芯片，玻片上的醛基可与 DNA 上的碱基 G、C、A 的芳香胺反应形成 Schiff 碱；还可以通过与氨基修饰的 DNA 反应形成 Schiff 碱而得以固定。如果用醛基修饰的玻片制备寡核苷酸芯片，则必须先在寡核苷酸的 5′端连接氨基碳臂后才能点样。

氨基修饰的玻片（例如多聚赖氨酸修饰的玻片）只能用于制备 cDNA 芯片，氨基在中

性条件下带正电荷，与 DNA 分子中带负电荷的磷酸基团通静电结合而被固定，经 UV 交联后氨基还可以进一步与 DNA 中胸腺嘧啶等碱基共价结合，以提高固定率。

从 Sigma 公司或 TeleChem 或其他公司都可以买到不同修饰表面的芯片。

（二）芯片阵列制作方式

（1）芯片阵列制作　下面介绍 4 种制作芯片阵列的方法（即点样的方法），前两种为合成后加样法，后两种为原位合成法。

① 点接触法　将纯化的、合成好的样品（如 cDNA 或寡核苷酸等）通过机械手直接点在载体上。

② 喷墨法　以定点供给的方式，通过压电晶体或其他推进形式从最小的喷嘴中将生物样品喷射到玻璃载体上，但喷嘴不与芯片直接接触。所加的样品需要事先合成并纯化。

③ 光刻原位合成法　将半导体工业中的光蚀刻术和 DNA 的化学合成方法相结合，在芯片点阵的原位上合成 DNA。也可用同样的方法制备光导向肽芯片，不同的氨基酸依次通过肽键连接和延伸。该法的主要缺点是需要花费大量的时间去设计和制造价格很高的照相掩蔽网。

④ 分子印章合成法　是一种在片压印半自动合成 DNA 芯片的装置，即采用分子印章，多次压印及原位定点合成 DNA 制备基因芯片。根据所需的微阵列设计制备有凸凹的微印章，在刻好的印章上涂上对应的单核苷酸，再按照设计的顺序将涂有不同的单核苷酸的微印章逐个依次压印在同一芯片上，然后根据清洗、脱保护、氧化及封闭等化学步骤进行原位 DNA 合成，直至达到所需的长度。

还有一些其他制作芯片阵列的方法，如通过导电的吡咯单体的聚合形成微阵列；或者将寡核苷酸固定在光纤纤维的一端后，再将光纤纤维捆绑成束，将在光纤纤维的末端形成由不同寡核苷酸的微阵列形成的集合体，这种制备方式仍然属于合成后加样的模式。

目前广泛采用的点样方式仍是合成后加样至玻璃片上的方式。

（2）芯片制备步骤简介　微阵列基因芯片的制备包括待固定样品的制备与纯化、点样、点样后的固定处理（例如紫外交联）、封闭（防止非特异性吸附）等步骤。在此以 cDNA 芯片的制备为例略作介绍。

① 用 PCR 扩增 cDNA 克隆（扩增片段长范围在 800～3000bp 之间），反应完成后，通过 Sephadex G-50 柱纯化除去未掺入的游离核苷酸。

② PCR 产物用异丙醇纯化后进行点样。DNA 可溶于水中（需要时可在水中加入 50% DMSO，防止点样时间过长，水分蒸发），也可以溶于 3×SSC 的盐溶液中。根据有些作者的经验，虽然用水溶解的样品在玻璃片上的固定率较高，但是均匀度、整齐度及面积大小均不如溶解在盐溶液中，建议还是使用盐溶液作为点样液。点样液的浓度一般为 0.5 $\mu g/\mu l$（$10\mu mol/L$ 左右）。点样时可根据实验室已有的条件，采用不同厂家生产的点样仪，手动点样仪或自动点样仪均可。

③ 点样后处理。将点样后的玻璃片进行扫描，以观察均匀度，然后经过水合、干燥等步骤处理。一般先水合 1～2h（置于湿滤纸上，60℃水浴中不同时间），再在室温下干燥 30min，然后置于紫外交联仪中进行交联（60mJ）。交联后再用 0.2% SDS 在室温下洗涤 5min，最后用 ddH_2O 冲洗 2 次，每次 1min，再进行扫描。也可在水洗后放入 0.2%硼氢化钠溶液中处理 10min，晾干备用。

点样时也可根据实际情况，自己摸索最佳条件。如果制备寡核苷酸芯片则还需要对上述条件再作些改进，但其大致过程和原理是相同的。

④ 芯片固定率的测定。为了评价 DNA 的固定效果，还需要检测 DNA 的固定率。固定率是指加样后的芯片经洗脱后，与玻片结合的 DNA 总量占总点样量的百分比。固定率 =（洗脱后的荧光值 / 洗脱前的荧光值）× 100%。

为此，可通过 PCR 技术将 Cy3 和 Cy5 标记（掺入）到一定长度的 DNA 样品上，然后将纯化后的 PCR 产物点于玻璃片上，每个样品点 20 个以上，用扫描仪及相应的软件测定和分析不同的加样条件及样品在洗脱前后的荧光强度，计算出样品固定率。

⑤ 用标记探针杂交实验进一步验证芯片的质量。根据前面实验摸索以确定实验的优化条件，在此条件下固定 cDNA 样品，并与人胎盘 mRNA 进行杂交实验，观察和检测表达谱芯片的信噪比，选择比值最高的 cDNA 芯片的制备条件为操作的标准化步骤。

（三）杂交探针的种类

基因芯片检测应用的是杂交原理，而核酸分子杂交则需要制备探针，以便直接检测靶基因或者检测外源病原体基因。为了提高检测的灵敏度，可采用两种措施：一是放大检测信号；二是扩增模板以增加模板量。对于后一种情况，可采用 PCR 扩增技术；而在此将重点介绍利用探针技术放大信号方面的方法和研究进展。

（1）分支探针　一般而言，这类探针分子具有两种功能区，即识别区和报告区，见下。

识别靶序列
- 靶序列短：杂交迅速
- 靶序列长：完成杂交费时长，甚至过夜

携带报告分子
- 携带越多，信号越强，灵敏度越高
- 可用核素或非核素方法标记报告分子

根据上述原理可设计出两类杂交法：①同时采用两种探针，一为捕捉探针，将其固定在载体上，另一探针的一半为识别序列，可识别并与靶分子结合，另一半则与捕捉探针互补结合，②直接将带识别靶序列的探针固定在载体上兼起捕捉作用。

探针除了有识别序列外，还需引入报告分子以供信号检测源检测。为此，可选用一个带有多官能基团的分支探针，每一个分支均与一个报告分子连接；分支探针的另一端则留出一个官能基团与载体上已固定的寡核苷酸探针（相当于上面提到的捕捉探针）相连。分支探针的使用使杂交信号得以成倍放大，见图 13-1。

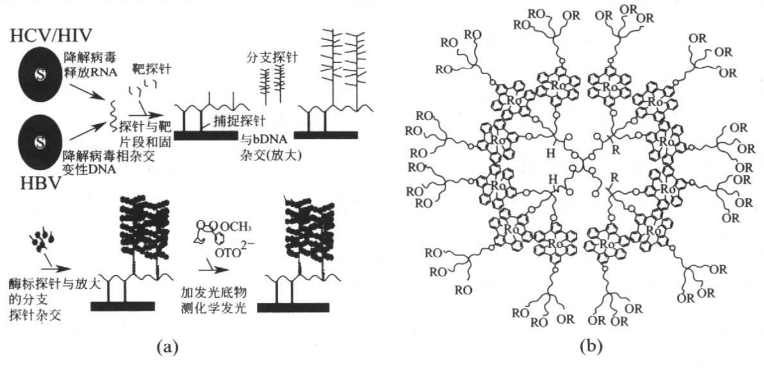


图 13-1　分支探针放大信号原理

（a）多探针夹心杂交原理；（b）含 12 个三联吡啶络合物的发光分子

（2）分子信标　这是一种设计巧妙的新型荧光标记核酸探针。在长度为15～30mer的序列特异性的寡核苷酸探针的3′端及5′端，分别加上5～8mer长度的互补序列，称为茎秆区。在自由状态时，由于茎秆区的存在，使探针分子形成发夹状结构，所以又称为发夹式探针；其形状又像海面上的灯塔，也有人称将这类探针称之为分子灯塔。

在分子信标的5′端及3′末端分别连接荧光素分子（F）及猝灭剂分子（Q）。在分子处于自由态时，发夹结构的存在将两末端拉在一起，从而使（F）与（Q）靠近。由于分子共振和能量转移，使F和Q之间的能量传递交换几乎达到100%，荧光被猝灭；当体系中由于靶分子的加入，分子信标中的环状区与靶分子的序列完全互补，形成异源双链杂交体时，信标茎秆区也随之被拉开，F与Q间的距离增大，信标分子的荧光几乎100%恢复。故分子信标在杂交前后的信号/背景比值差别极大，见图13-2。

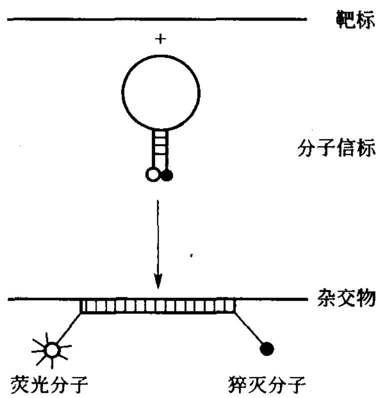


图 13-2　分子信标工作原理示意图
黑色的点代表猝灭剂分子（Q），
空心的圈代表荧光素分子（F）

综上所述，信标分子一般包括3个部分。

① 识别部分　是与靶分子杂交的部分，长度一般为15～30mer。若在杂交双链中有一个错酸碱基，其热稳定性显著下降，故分子信标可用于检测点突变。

② 茎秆部分　位于信标分子两端，长度为5～8个序列互补的核苷酸。

③ 荧光素、猝灭剂和连接臂，荧光素连接在信标分子的5′端，常用DABSY（二甲氨基偶氮苯甲酰基）荧光染料。荧光素、猝灭剂与核苷酸之间的连接，需要分别引入连接臂，如$(CH_2)_6$-*S*-CH_2-O及$(CH_2)_7$-*N*-两种类型的碳臂。

信标分子能在均相系统中实时测定出靶标的含量，而不必分离去除多余的信标分子。其主要应用如下：①实时定量PCR测定靶标浓度；②PCR扩增并进行点突变的检测；③多色荧光分子信标检测人的等位基因；④用分子信标PCR法目测测定结核杆菌的耐药性。

（3）肽核酸（peptide nucleic acid）探针　肽核酸是以中性酰胺键为骨架的一类新的寡核苷酸的类似物，故称PNA。PNA与寡核苷酸的磷酸二酯键骨架不同的是它以甘氨酸结构单元为骨架；PNA中的碱基部分则通过亚甲基羰基与主骨架相连，碱基与骨架间隔3个化学键，相邻碱基之间相隔6个化学键。PNA与天然寡核苷酸的结构见图13-3。

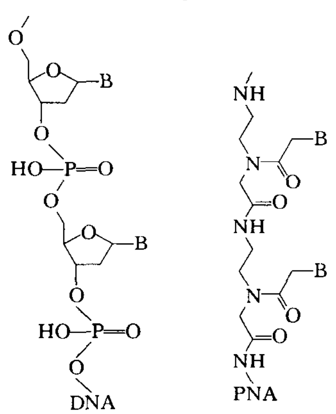


图 13-3　寡核苷酸和肽核酸的结构图

二者的比较结果为：①与天然核苷酸结构相同之处是它们的碱基均能通过Watson-Crick氢键与互补DNA的特定核酸序列配对结合；②由于PNA分子不带电荷，互相之间及与核酸杂交时均不存在静电排斥作用，用它代替寡核苷酸探针固定在芯片载体上，对杂交反应具有更大的优越性；又由于PNA分子缺乏核酸分子所具有5′→3′的极性，因此杂交时可以正向或/和反向方式与靶序列互补结合（以反向结合占优势）；③PNA由于缺乏磷酸二酯键，能耐受核酸酶的降解，比较稳定；④PNA也可以用固相肽合成技术合成，制备方便；⑤PNA与靶序列杂交时，对碱基错配有

很强的识别能力，例如一个 15mer 的 PNA/DNA 异源双链中，有一碱基错配时，T_m 值将下降 8～20℃（平均 15℃），而对应的 DNA/DNA 复合物中，一个碱基错配时，T_m 值只下降 4～16℃（平均 11℃），这种高度的区分能力可以建立以 PNA 为基础的分析点突变的方法；⑥PNA 能识别 DNA 双链上的特定碱基序列，形成局部的 PNA/DNA/DNA 或 PNA/DNA/PNA 三链复合物。

（四）样品制备和标记

主要指待检测的靶基因片段的体外扩增和标记。

在生物芯片中，待测样品还需要经过体外扩增并加以标记才能使靶基因杂交后有足够的信号强度供检测用，特别是采用诊断芯片检测血液中微量病原微生物的基因组核酸时，杂交前一定要对样品进行扩增，并在扩增的过程中标记靶 DNA。靶基因体外扩增的方法很多，但常用的仍然是 PCR 方法。

样品的标记物主要还是采用荧光素标记。标记的方法：①可以利用荧光标记的引物，进行 PCR 扩增反应；②也可以使用荧光标记的一种三磷酸脱氧核糖核苷酸（dNTP），如 Cy3-dCTP/UTP、Cy5-dCTP/UTP，使它们在 PCR 扩增的过程中掺入到靶基因序列中。目前常使用的荧光物质有荧光素、罗丹明、HEX、TMR、FAM、Cy3、Cy5 等。

PCR 扩增后的产物分离的方法也很多。可以用 PAGE 分离、亲和素偶联磁珠（PCR 扩增时，一条引物用生物素标记，另一条用荧光素标记）捕捉扩增产物等。

样品制备和标记的过程大致如下。

① 样品总 RNA 的提取。用 Gibco 公司的 Trizol 试剂一步法提取细胞或组织的总 RNA。详见说明书，主要步骤如下：细胞用冰冷的等渗 PBS（pH7.2）洗涤 2 次，加 Trizol 试剂裂解细胞，氯仿抽提一次，离心后吸出上层液相，异丙醇沉淀 RNA，75%乙醇洗 2 次，空气中干燥，溶于缓冲液中，用紫外分光光度计测定 OD_{260nm}/OD_{280nm} 的比值，并用甲醛凝胶电泳观察 28S、18S、5S 条带的比例，以确定总 RNA 的提取质量。一般还需要测定 OD_{260nm}、OD_{280nm}、OD_{230nm}、OD_{320nm} 的值以及 OD_{260nm}/OD_{280nm} 比值。

② mRNA 的分离和纯化。用 mRNA 纯化试剂盒（例如 Qiagen 公司的 Oligotex mRNA Midi Kit 或其他公司的试剂盒），操作步骤可见说明书。

③ 逆转录生成 cDNA。参照 Schena 的方法逆转录，标记 cDNA 并纯化。用 Cy3-dUTP 标记对照细胞的 mRNA，用 Cy5-dUTP 标记实验细胞的 mRNA。逆转录反应完成后，乙醇沉淀并溶于适量 5×SSC+0.2%SDS（随条件不同，也可用 0.1%SDS）杂交液中。

（五）杂交

将基因芯片和标记的杂交探针分别在 95℃水浴中变性 5min，将探针加在基因芯片上，用盖玻片封片，杂交面积由所用的盖玻片的面积决定，一般每平方厘米需要杂交液 2mL。置 60℃或 62℃杂交 10～17h，然后揭开盖玻片，分别用 2×SSC+0.2%SDS 、0.1%×SSC+0.2%SDS 、0.1%×SSC 洗涤 10min。移去盖玻片时，不可用力掰下盖玻片，可把玻片垂直放在 0.1% SDS、0.3×SSC 液体中，来回晃荡，使液体浸润盖玻片，让盖玻片自由滑下。在清洗时玻片要水平放置，以使背景均匀。清洗时，可在摇床上摇晃。玻片清洗后室温晾干或放在 50ml 的锥形离心管中，2000r/min 离心 1min，以除去玻片表面的液体。此时即可用双通道激光扫描仪进行扫描。

常规杂交反应由于受到探针解链温度、探针长度和溶液中靶序列的初始浓度的影响，必须选择最理想的条件以尽可能使正确配对的序列不被遗漏，并使杂交错配降到最低。

探针必须具有相似的解链温度，这样才能保证杂交条件的一致性，这将取决于探针的组成和长度；即便探针的长度相同，由于GC含量的差异也具有不同的解链温度，导致杂交动力学的差别。

杂交后的DNA芯片在严谨条件下洗涤以去除未杂交的残留核酸样品。

杂交反应与溶液中靶序列的初始浓度成正比，所以必须提高靶DNA浓度以加快反应速度。

（六）杂交信号采集和处理

用不同公司的扫描仪及配套软件分析Cy3和Cy5两种荧光信号的强度和比值并确定其可信限。芯片扫描中的关键问题：①芯片保持洁净，由于芯片扫描仪有极高的灵敏度和分辨率，污染物会引起噪声和背景，故在芯片制作杂交和清洗过程中应注意恒温和恒湿，并在超净台（或罩）或超净室中进行，所用溶液和试剂的纯度应符合规定和过滤除菌；②细胞或组织样品测定时会产生较高的背景噪声；③芯片完成杂交和清洗反应后应立即测定，防止靶分子降解和荧光猝灭；④芯片扫描仪应放置平稳，防止外源性震动。

一般基因芯片杂交信号种类有：①荧光标记（单标或双标）的靶DNA（FITC、Cy3、Cy5等）检测，如FITC/Cy3、FITC/Cy5、Cy3/Cy5；②酶标的靶DNA（灰度图像分析）；③其他种类的靶分子标记（单彩色或多彩色），如化学发光、放射性核素标记、电化学等。

不同的杂交信号，检测扫描装置也有所不同。目前，商业化的芯片扫描仪主要有两类：激光共聚焦芯片扫描仪和CCD芯片扫描仪。此外，还有用荧光显微镜自己组装搭建的DIY扫描装置。

1. 激光共聚焦芯片扫描仪

生物芯片中，靶分子常用2种或2种以上的荧光染料标记，故扫描结果常用两种荧光光密度比值来表示，以减少或消除测定时的干扰和实验误差，提高结果的可靠性。为此，商业化的扫描仪常采用2种以上不同波长的激光器作为激发光源。

扫描动作由马达驱动的芯片平台作X-Y二维移动来实现。扫描结果由计算机采集及处理。扫描得到的图像是黑白的。多种荧光标记的芯片，可分别在不同波长的激光下重复扫描进行；也可在多个激发波长下，同时扫描（即一次扫描可以得多个波长的数据）。后者可克服多次扫描时由于扫描平台机械性能重复性不好造成的误差。

2. CCD芯片扫描仪

这种扫描仪结构较简单，激发光源多采用氙灯或高压汞灯，单色系统采用激发窄带干涉滤光片，与CCD（charge coupled device）摄像头相连。由于CCD芯片扫描仪是同时读取整个芯片结果，不需要扫描平台的二维移动，扫描速度快于激光聚焦扫描仪。其工作原理为：光源→激发窄带干涉滤光片→芯片→发射窄带干涉滤光片→透镜→CCD摄像头→CCD芯片成像→图像转换→数字信号→计算机存储。

（七）生物信息的分析和提取

芯片扫描数据的采集、处理、分析和报告是生物芯片检测的重要环节。一张汇集了数十个到数千个点识别分子的芯片，它们杂交后的光密度值必须及时收集和处理；如果芯片采用多色荧光染料标记，还必须具备和处理各种荧光强度的比例信息，因此，必须有一个专门系统来处理芯片的数据，完整的数据系统应由下列几步骤组成，分述如下。

1. 图像分析和数据提取

扫描得到的图像是数字文件，而不是各样品点的OD值、面积、OD比值等数据，必须

经过图像处理提取各样品点的数据，供进一步的生物学统计分析。

图像分析软件功能如下。

（1）图像的平滑过滤　去除芯片制作、杂交、清洗和测定过程中的污染、细胞和组织碎片、芯片扫描仪的噪音等所引起的干扰信号。

（2）图像背景的确定　图像上各点的光密度值包含了样品的信号和背景的信号，在提取各样品数据前必须将背景扣除。

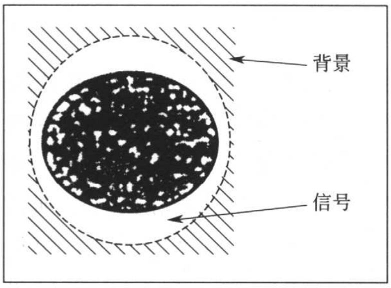


图 13-4　样品斑点的识别示意图

（3）样品斑点的识别　一般是在图像中选择需识别的区域，输入阵列的行数和列数后，计算机便自动产生一个网格框，套在芯片图像上的每一个样品点上。计算机在框内中心再设置一个圆，使样品点完全落在圆内。圆内信号即为样品信号，圆外信号为背景信号，背景信号和样品信号之间留出一圈过渡缓冲区（图 13-4）。

对双色荧光标记的芯片，在图像分析前，必须将 2 个图像进行重叠，使两个图像有相同大小的网格和内圆，并定义其中一个为参比图像。

（4）数据的提取、存储与显示　生物芯片中各点阵提取的数据包括 OD 积分值（样品斑内各像素的光密度值之和）、平均光密度值（积分值/样品斑点内像素总数）、样品斑点面积（斑点内像素总数）和样品斑点内光密度最大值和最小值。如果是双色荧光标记的芯片，还需提取各个样品点中两种不同荧光的光密度比值。在由阵列图像转变为数据阵列后，数据提取的任务便完成了，可以储存或输入到特定的数据库中保存。

为直观地观察数据，大多数商业软件都采用伪彩图、散点图和直方图来显示数据。

伪彩图可直观地观察芯片的阵列图像，最强的一般用白色，最弱的黑色，中间用其他颜色如红、橙、黄、绿、蓝等表示，见图13-5。

散点图的横坐标表示参比图像的光密度或其对数值；纵坐标表示非参比图像的光密度或其对数值；过原点的 45°线则表示两种荧光标记探针的信号强度相等。用鼠标任击其中一点均可显示出该点对应的 2 种荧光标记的图像的光密度值、光密度值标准差及表达百分比。见图 13-5。

直方图的横坐标表示图像的光密度值；纵坐标表示具有该光密度值的像素的个数。

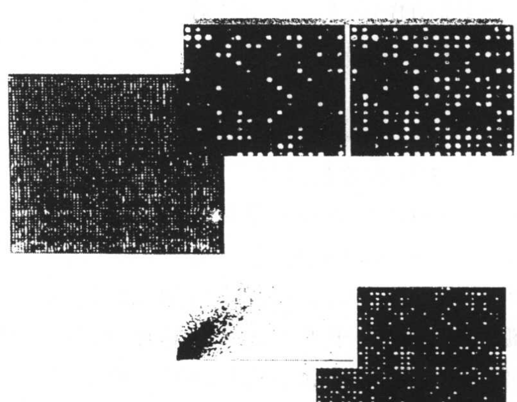


图 13-5　散点图与伪彩图对照

2. 芯片数据信息的管理

数据库管理的目的是方便地保存、管理和查询数据信息。管理系统应保存下列信息：芯片各点的基因名称、基因的碱基序列或功能的主要描述、Genbank 保存取码、IMAGE 克隆标识、代谢途径标识、内部克隆标识，芯片各点的荧光图像、背景值、光密度值，光密度比值等；还应有芯片的制备、实验、环境条件、样品的类型和制备方法、实验目的和操作者姓

名等信息。

3. 芯片数据的分析

通过芯片各点数据的分析比较或芯片间的数据比较来实现芯片数据的分析。DNA 芯片能同时平行分析大量的信息，包括 SNP、表达序列标签（EST）和基因克隆、其他有关细胞的基因表达等。

数据分析主要包括以下内容。

（1）芯片数据的归一化　由于直接比较多个芯片表达的结果是比较困难的，必须减少或消除各个芯片实验之间的差异。常采用下述方法。

① 用“看家基因”作为特定的对照基因。杂交时将一定量的与其互补的荧光标记探针混合到杂交液中，这样就可以将对照信号与各样点信号比较，该比值便可消除各实验点间的差异，从而达到归一化的目的。

② 将各点光密度值或比值除以所有点的平均值；或附带一些参数如平均值等以作为该芯片的内部对照。

除上述的归一法处理外，为比较多个芯片的数据，还应严格控制每次的实验条件，如靶 DNA 标记的程序、荧光激发和发射的效率、测定的条件等，使实验在相同的环境和条件下进行。

（2）芯片数据的视图分析　是一种最简单、最直接、最直观的分析方法。通常用散点图（二维或三维）、直方图和饼图直观地显示芯片表达的结果。

（3）芯片数据的统计分析　目前常用聚类分析，这是一种在事物面孔尚不清楚的情况下研究生物的分类方法。它直接比较样本中各指标之间的性质，将性质相近的归为一类，性质差别较大的归在另一类，可帮助研究人员发现新的基因、新的 DNA 序列、基因的突变位点等。聚类分析根据聚类指标或计算机方法的不同又可分为许多种方法。

（4）芯片的生物学分析　根据视图分析和统计分析的结果，结合生物学知识对实验结果作出相关判断。例如，时间过程分析是采用较多的一种分析方法，可以得到不同发育时期的基因表达、用药前后基因表达的变化等有用的资料。

第三节　蛋白质芯片

21 世纪出现的新型生物芯片——蛋白质芯片（protein chip）和基因芯片（gene chip，DNA chip，DNA microarray）一样，同属于生物芯片的范畴，蛋白质芯片比基因芯片更接近生命活动的物质层面，因而有着比基因芯片更加直接的应用前景。蛋白质是一切生命活动的基础，受基因表达的调控，而蛋白质本身亦具有特有的活动规律，蛋白质在合成后具有相对独立的修饰、运转和相互间作用能力，同时还具有对外界因素发生反应的能力。基因的主要功能是通过其表达产物——蛋白质来实现的，在转录时一个基因可能以多种 mRNA 形式剪接，并且同一蛋白质可能以多种形式进行翻译后的修饰。因此，一个蛋白质组不是一个基因组的直接产物，蛋白质组中蛋白质的数目可以超过基因组的数目。蛋白质表达谱比 mRNA 表达谱更能真实地反映生物体的功能机制。人类基因组大规模测序工作已经完成，随着

以功能基因组学和蛋白质组学为主要研究内容的后基因组时代来临，蛋白质芯片作为检测蛋白质存在和运动变化的高效工具，必将发挥越来越大的作用。

一、基本原理

1. 蛋白质芯片的概念

蛋白质芯片（protein chip）的狭义的定义是指用专门用于蛋白质功能研究及相互作用分析的蛋白质微阵列（protein microarray）或一类生物芯片，即采用原位合成、机械点样或共价结合的等方法将多肽、蛋白、酶、抗原、抗体固定于硅片、玻璃片、塑料片、凝胶、尼龙膜等固相介质上形成的生物分子点阵，见图 13-6。在待分析样品中的生物分子与蛋白质芯片的探针分子发生杂交或相互作用后，利用激光共聚焦显微扫描仪对杂交信号进行高通量检测和分析；它是将整个蛋白质水平的相关生化分析过程集成于芯片表面，从而实现对多肽、蛋白质及其他生物成分进行高通量检测。

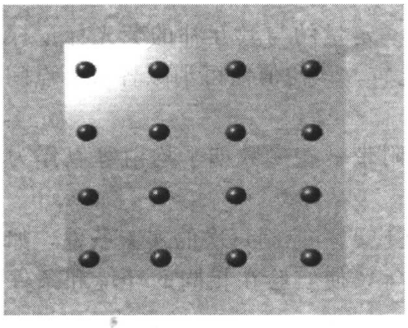


图 13-6　蛋白质芯片

和基因芯片相比蛋白质芯片有自身的特点。一般而言，基因芯片只能检测到人体各种遗传信息在 DNA/RNA 水平的变化，而无法检测出在蛋白质水平上的变化。同时基因芯片的制作和检测比较费时、所需探针必须人工合成、必须经 PCR 扩增过程、还要进行昂贵的荧光标记、杂交结果还存在批间误差、检测结果的分析较复杂。蛋白质芯片的反应结果表明，只要抗原与抗体（或受体与配体等）的结合是特异性的，在同一体系中相互之间就不会发生影响，可进行并行反应。有人做过这样的实验，选用两种不同来源的抗原分布于同一张 PVDF 膜上，同时还将它们所对应的抗体混合于同一体系中，实验结果较少受到干扰。由此可以说明，如果将多种蛋白质（抗原或抗体）分布于同一张 PVDF 膜上，有可能对同一份血清标本，同时一次检测不同的抗原（或抗体）。这样，取一份患者的血清，就可同时测得多个指标，在临床疾病检测和诊断上有较好的应用。

在此，举一个例子说明在疾病诊断方面基因芯片和蛋白质芯片各自的特点和适用范围。

如果要得知患者得的是甲肝、乙肝或丙肝等肝炎中的哪一种，常用肝炎检测基因芯片来诊断。首先将病人的血清提取物分别作 PCR 扩增，将分别针对甲肝、乙肝、或丙肝的病毒 RNA 或 DNA 片段对应的引物、dNTP（在其中任意一种碱基上标上荧光标记）、聚合酶、血清提取物分别混合后，进行 PCR 扩增，这些扩增物中必定有一种中含有扩增片段。PCR 产物解链为单链后，将这些扩增出的物质混合，与检测芯片上的探针进行杂交，根据带有荧光标记的扩增片段是否与探针片段互补结合，就可以判断病人患有哪种肝炎。总之，利用基因芯片来检测，操作复杂，须查阅资料合成引物，须经 PCR 扩增和荧光标记，价格昂贵，试剂材料消耗大，耗费大量人力与时间。

由于患者血清中的肝炎病毒可视为一种抗原，所以将各型肝炎的抗体分别固定于载体上，制成肝炎检测的蛋白质芯片。只需在该芯片中加入患者的血清，保温 1h 后，再加入酶标二抗，反应显色，如血清中含有抗原，就能很快就能知道患者得的是何种肝炎。整个过程所需时间较短，操作并不复杂，所需材料用量较少。通过以上肝炎检测基因芯片与肝炎检测蛋白质芯片的比较，可以看到蛋白质芯片有其独到之处。

此外，基因芯片是对基因进行 DNA 或 RNA 水平上的检测，但是只有当致病基因表达出蛋白质后才表现为疾病，从某种角度来说，蛋白质芯片的检测比基因芯片的检测更重要。

随着人类基因组计划测序工作的完成，在后基因组时代，将要进入更加复杂的蛋白质功能的研究，迫切需要高通量的蛋白质芯片技术。

2. 蛋白质芯片的特点

① 高通量，高效率。

② 蛋白质芯片是一个动态的技术流程概念，不局限于某个固定的形态或结构。一般而言，蛋白质芯片有两种加样方式：一种是在芯片上排布蛋白质微阵列；另一种则是将特异性配体置于芯片上进行点阵，用以筛选和检测复杂生物体系（如细胞提取液等）中的靶蛋白。

③ 适用于包括组织、细胞系、体液在内的多种生物样品中特殊蛋白质靶标的鉴定和筛选。芯片上固定的样品包含针对信号传导、癌症、细胞周期调控、细胞结构、凋亡和神经膜电位等广泛的生物功能的相关蛋白，跨度大、适用范围广；如果蛋白质芯片技术与基因克隆表达、蛋白组学的 2D 及 MS 技术结合起来，进一步对芯片上筛选出来的已知或未知功能蛋白质进行鉴定，对发现新的疾病相关基因或蛋白十分有意义。

④ 芯片上进行的结合或杂交反应的产物可以用不同方法检测，从最简单的 ELISA 方法到荧光 CCD 成像、放射自显影等，灵敏度高达 pg/ml；某些开放性的芯片平台设计，可以用各种型号 DNA 芯片荧光扫描仪进行检测。

⑤ 在生命科学及医学领域中的应用前景广泛。

3. 蛋白质芯片的分类

按照不同的标准可对蛋白质芯片进行不同的分类。

(1) 按芯片表面样品结合方式分类　分为化学型和生物化学型。化学型蛋白质芯片表面包括经典的色谱分析表面，如结合普通蛋白质的正相表面、用于反相捕获的疏水表面、阴阳离子交换表面和捕获金属结合蛋白的静态金属亲和捕获表面。通过介质的疏水力、静电力、共价键等结合样品中的蛋白质，然后经特定的洗脱液除去杂质蛋白质，而保留感兴趣的蛋白质。这种芯片特异性较差。

生物化学型蛋白质芯片则是把生物活性分子（如抗体、受体、配体等）通过生化反应共价结合于预先活化的表面阵列，用于捕获样品中的靶蛋白。由于生物化学型芯片具有高度的特异性及生物活性分子的多样性，其应用范围和前景明显优于化学型蛋白质芯片。

(2) 按点样蛋白质有无活性功能分类　分为无活性的（non-living）和有活性（living）的蛋白芯片。无活性的芯片是将已经合成好（或纯化好）的蛋白质点在芯片上，有活性的芯片则是在芯片上点上生物体（如细菌），在芯片上原位表达蛋白质。

无活性的蛋白质芯片制作方式主要分三类。①原位合成。Ceysen HM 等采用荧光甲氧羰酰氨基酸保护策略，由 C 端向 N 端，共合成了 208 个六肽，分别排列在直径为 4mm 的聚乙烯膜上，用以研究一个由 213 个氨基酸残基组成的蛋白质的结构活性域。②点合成。A. Krame 等利用该法在纤维素膜上合成多肽文库，纤维素膜上有丰富的羟基，对于多肽的合成和固定十分有利。③光蚀刻术。类似与基因芯片的原位合成法，采用光蚀刻术固相合成的多肽芯片密度为每平方厘米 25 万点，多肽的平均长度为 5 肽，用这种芯片成功地进行了抗体检测。

活性芯片可以提供模拟的体内环境，对于蛋白质功能分析更为有利。P. Uetz 等利用酵母细胞构建了第一个全基因组的活性蛋白质芯片。活性蛋白质芯片通过一次实验就能得到上千种

未知蛋白质的功能线索，很快就可以查明酵母细胞的某一开放阅读框架（ORF）是否表达为一种分泌蛋白抑或是一种胞内酶。此外，还可以在芯片上进行活细胞的遗传性状筛选。

（3）按蛋白微阵列的种类和检测目的分类　分为蛋白质表达芯片（protein expression chip，PEC）和蛋白质功能芯片（protein function chip，PFC）。PEC 能观察到不同条件下哪种蛋白质被合成而获得基因表达的详细信息；PFC 则能分析众多蛋白质的功能性质。PEC 的目标与 DNA 芯片相类似，它将大量检测用的分子（常为抗体）以微阵列方式固定在芯片表面上，故加样反应后就能查明样品中有无这些分子的靶标。目前，虽尚不能在芯片上固定包含机体全部蛋白质的文库。这类芯片（即便只含有蛋白质的亚群）在生物学研究和新药开发中也有着重要的使用价值，特别是将传统的蛋白质功能分析和以核酸为基础的基因表达分析技术联合应用时，它的重要性更突出。

PFC 与 PEC 有许多共同特点，即它们都需要有大量不同分子固定在芯片表面；不同的是 PFC 表面上固定的却是待测的蛋白质分子，随后再用其他已知配体、药物和生化药品来筛选各待测分子与它们结合的情况，并了解其结合的特异性及分子间的相互作用力。

（4）按检测方法分类　可分为多种，包括 BIACORE 的单元芯片、多元蛋白质光学芯片和 SELDI 质谱芯片等。

这类芯片的共同特点都是将生物分子作为配基，以不同的排列形式固定在固相芯片表面或表面微单元上，利用生物分子间的特异结合的自然属性，待测分子与配基分子在芯片表面会形成生物分子复合物。然后，检测此复合物的存在与否，达到对蛋白质的探测、识别和纯化的目的。以上不同技术的差异仅在探测方法的不同。BIACORE 技术利用表面等离子体共振技术扫描芯片，进行单一蛋白质检测；多元蛋白质光学芯片是利用光学成像法原理，可同时检测多种混合的蛋白质；SELDI 技术则采用质谱法，以时间顺序检测序列蛋白质。

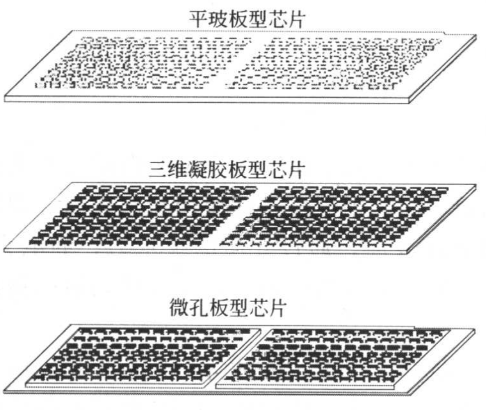


图 13-7　不同板型的蛋白质芯片

（5）按固相载体片基分类　分为平玻板型（plain-glass slide）、三维凝胶板型（3D gel pad chip）和微孔板型（microwell chip），见图 13-7 及表 13-1。

表 13-1　不同类型蛋白质芯片的性质比较

平玻板	三维凝胶模板	微孔板
可与标准微阵列加样仪及探测设备兼容	可与标准微阵列加样仪及探测设备兼容	可与标准微阵列仪及探测设备兼容，但需要联合使用
样品高挥发	样品挥发减少	样品挥发减少
对复合反应效果弱	可进行液相反应，但对材料需较长洗脱时间	广泛用于液相及复合反应
成本低	成本高，需光制板技术	成本低
可能会有交叉污染	无交叉污染	无交叉污染

（6）按工作原理分类　可分为探针型和凝胶电泳型。探针型类似于 DNA 芯片，即在固相支持物表面高密度排列的探针蛋白点阵，可特异地捕获样品中的靶蛋白，然后通过检测器

对靶蛋白进行定性或定量分析。另一种就是微型化的凝胶电泳板。在电场作用下，样品中的蛋白质通过芯片上的分子筛微孔彼此分离，经喷雾直接进入质谱仪进行检测，以确定样品中蛋白质的分子量及种类。

二、蛋白质芯片制备和检测的基本流程

对一个理想的蛋白质芯片的初步模型和检测过程应该是：选择一种能够牢固结合蛋白质分子（抗原或抗体）的固相载体，这样形成蛋白质的微阵列，即蛋白质芯片；加入与之特异性反应的带有特殊标记的蛋白质分子（抗体或抗原），两者结合后，通过对标记物的检测来实现抗原抗体的互检，即蛋白质相互反应的检测。此种方法构建的模型所需蛋白质的量极少，反应相对较快，蛋白质芯片稳定性较好，灵敏度较高。

（一）蛋白质芯片制备与检测流程

总的流程图可见图 13-8。

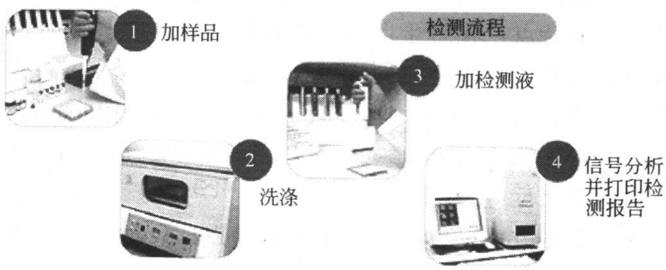


图 13-8　蛋白质芯片检测流程

1. 制备抗体文库

如利用噬菌体展示技术。

2. 选择固相载体片基

呈薄片状，外观可做成长条状、圆形或椭圆形等不同形状，经特定化学修饰处理后能承载和吸附相关的生物样品制剂。常用的材质有硅、云母、镀镍或镀金的各种膜片等。各公司制作芯片采用的载体材料基本相差不大。

有的作者为了验证不同固相载体片基对蛋白质分子的固定效果，选取了 7 种材料进行实验，包括普通玻片、氨基修饰的玻片、醛基修饰的玻片、聚赖氨酸修饰的玻片、聚苯乙烯板（酶标板）、醋酸纤维薄膜（NC 纸）、特殊处理的 PVDF 膜（95%酒精浸泡，蒸馏水洗涤，烘干，备用）。实验结果表明 PVDF 膜的固定效果最佳，经过多次洗涤，PVDF 膜上的显色反应斑点仍很清晰。与其他载体相比，PVDF 膜的优点在于：①固定蛋白质的能力较强，通过 9 次含表面活性剂的洗涤液洗涤，仍能牢固结合蛋白质分子，使在其上进行相关的反应成为可能；②因 PVDF 膜为纸质，能够被剪裁成所需的任何大小，使用较方便；③PVDF 膜作为进行免疫酶联反应的载体有良好的效果，各种不同来源的抗原都能在 PVDF 膜上与其相应的抗体反应，并呈现较清晰的浓度梯度。

由于不管是那一类芯片都是将原来在溶液中进行的反应转移到膜上进行，不可避免地会造成“不适应”的结果，特别是蛋白质的功能活性与水环境密切相关。2004 年，Hamachi 及合作者介绍了一种新型水凝胶材料可用作富水性芯片片基，蛋白质反应和检测均可在其上完成。这种新型水凝胶的特点是它同时拥有亲水性和憎水性的纳米级间隔；前者能够捕捉大

量水分保证蛋白质活性，后者则能够放大荧光信号，以利检测蛋白质活性。

3. 制备探针蛋白质

对蛋白质芯片来说，探针的种类包括特定蛋白质的抗体或受体、结合某些阳离子或阴离子的化学基团、吸水或疏水物质、酶、免疫复合物等，通过化学结合反应固定在载体表面上，形成一个个点状芯池。

单克隆抗体具有高度的特异性和亲和性，是一种比较好的探针蛋白，用其制作的抗体芯片主要用于检测蛋白质表达丰度及确定新的蛋白质。还可以通过 2D 电泳分离技术分离得到的蛋白质作为抗原，把固相化的抗原和抗体库孵育，即可得相应的单克隆抗体。

噬菌体抗体库技术是典型的探针蛋白文库制备技术，其他的蛋白质文库，如全合成人重组抗体库、噬菌体肽库、噬菌体表达文库等也可用来制备探针蛋白。利用机器人将抗体文库以阵列形式固定在芯片载体上。

4. 蛋白质的固相化（点样）

由于蛋白质具有易于变性、不易于与固体表面结合、在固相支持物表面上不易折叠成有功能的构象而丧失其活性的特点，在将蛋白质固定在固体支持物表面时，必须保持蛋白质的立体构象，才能使蛋白质探针具有生物活性。

最简单的点样方式可采用手工点样，在一片 2cm×3cm 的 PVDF 膜上，手工点样可排列 100 个左右的点，每一点体积仅有 100～200nl，用量极少。加样完毕后，一般再固定 2h 以上，才可进行下一步的杂交反应。

此外，和基因芯片的加样方式一样，还可采用自动点样仪加样，有的自动点样仪在一个 servo 控制的 3 轴线性驱动系统上携带 384 个点样头，定位可以准确到 5μm，产生的密度大约为 300 点/cm^2。样品吸头的直径取决于点样的需要，可以在 150～450μm 之间调节变化。

采用直接点样法可以避免蛋白质的空间结构改变，保持探针蛋白和样品的特异性结合能力，但是如何将蛋白质固定在芯片片基上而不脱落，也是一个关键问题。生物化学家 MacBeath G 和化学家 Schreiber S 联手解决使蛋白质黏附于固相载玻片上而保持活性这一关键技术突破之后，才形成真正的蛋白质分子微阵列，使印迹蛋白微点阵用于高通量功能研究成为现实。技术的关键突破在于对载玻片的处理。用含醛的硅烷试剂处理玻片，并用含有 30%甘油的 PBS 溶解蛋白质后，再用 Affimetrix 生产的 GMS417 阵列点样机将微量蛋白质点在玻片上，这样，蛋白质的伯氨基与醛形成希夫碱（Schiff 碱）而将蛋白质以共价结合的方式附着在玻璃表面，再用 BSA 处理载玻片表面，然后用 *N*,*N*′-二琥珀酰胺基碳酸盐溶液（*N*,*N*′-disuccinimidylcarbonate）活化 BSA 的赖氨酸、谷氨酸残基，使其易与固相蛋白的表面氨基作用形成共价脲连接或酰胺连接，被固定的肽或小分子蛋白则可在载玻片 BSA 单层分子表面保持立体构象，易于与溶液中的大分子起反应，这样就可以与蛋白质进行杂交了。

和 DNA 芯片不同，蛋白质芯片很难制备的另一个原因是因为蛋白质离不开水，因此常规的干芯片很难满足要求。Arenkov 等把探针蛋白固定于微型聚丙烯酰胺凝胶块中构建了蛋白质芯片。由于聚丙烯酰胺凝胶拥有亲水性，能够捕捉大量水分保证蛋白质活性，同时凝胶的疏水性有利于荧光信号放大及检测蛋白质活性。微电泳以后，又能加速蛋白质的扩散及捕获样品中能与探针蛋白特异结合的靶蛋白，这类芯片很值得推介。

A. Lueking 等首先将人胚脑 cDNA 表达文库克隆至载体上，再用 IPTG 诱导原位表达包含 $(His)_6$ 表位（His-tag）的融合蛋白质。该步骤在 96 孔的微量滴定板上进行。接着将 96

孔的微量滴定板中的呈溶解状态的蛋白质点在 PVDF（polyvinylidene-difluoride）膜上。PVDF 膜具有极高的蛋白质结合力和机械强度。他们利用的是一种自动化转运印章技术，可以进行三维定位，转运印章设计的点样密度上限是在一张 222mm×222mm 大小的膜上排 10 万个点，在 $1cm^2$ 上最多能点上 300 个样品。1 个转运印章包含了 16 个点样头，其中 4 个点的是对照，另外 12 个点的是 6 种样品，一式两份。蛋白质点在 PVDF 膜上之后，蛋白质芯片就制成了。

现已有预置好已知探针的多种蛋白质芯片成品出售，如微阵列式、微孔板式、凝胶块状等类型。一般来说，一块载体上根据需要固定的芯池可多可少，例如作为研究用芯片的芯池数目较少，大多都在 6～10 个左右，多者可达 10 多个。但用于规模生产用的芯片，其芯池数量要多得多，有一种构象型蛋白质芯片，其芯池数目高达 1600 点/cm^2，呈微阵列排布；点样时须用机械手进行，可同时检测数千个样品。使用时先将需要检测的含有蛋白质的标本如尿液、血清、精液、组织提取物等按一定程序做好层析、电泳、色谱等前处理，然后在每个芯池内点入需要的样品，一般样品量只要 2～10μl 即可。根据检测目的不同，可选用含有不同探针的芯片。让标本在每个芯池中与特定的探针结合或与其中含有的生物制剂相互作用一段时间，然后洗去未结合的或多余的物质，将样品固定等待检测即可。

（二）待检样品的制备和标记

1. 常用样品

样品可以是酶制品（如碱性蛋白酶等）、细菌或细胞提取液（如阿米巴提取液）、抗原或抗体（如乙肝表面抗原/抗体、重组 e 抗原/抗体等）、生物活性物质（如甲状腺素等）、病人或正常人体液标本（如血清、尿、脑脊液、组织液、关节腔滑液、支气管洗脱液、细胞裂解液和各种分泌物等）。

2. 分级分离技术制备样品

样品（蛋白质、多肽、酶等）制备通常可采用分级分离技术将细胞或组织中的总蛋白质分成几部分，使低表达的蛋白质得到富集。主要预分离的方法包括电泳或色谱预先分离、蛋白质亲和富集、细胞器分级分离等。重点介绍以下几项技术：① 液相等电聚焦，可以将复杂的蛋白质混合物在一定的 pH 范围分离，将其浓缩在相对较窄的范围内；②2D 电泳，利用蛋白质的等电点和分子量通过双向凝胶电泳的方法将各种蛋白质区分开来是一种很有效的手段；目前，二维色谱（2D-LC）、二维毛细管电泳（2D-CE）、液相色谱-毛细管电泳（LC-CE）等新型分离技术都有补充和取代双向凝胶电泳之势；③吸附色谱，常用的有离子交换色谱（IEC）、疏水相互作用色谱、反相色谱和亲和色谱、体积排阻色谱（SEC）等，色谱与 2D 电泳联合分离可用于大量蛋白质的快速纯化；④亚细胞器分离，可对特定的蛋白质组分进行分离纯化。

最近发展的在组织水平上分离和制备蛋白质组样品的方法，如采用激光捕获显微切割技术（laser capture microd issection，LCM）方法分离癌变上皮细胞。其基本方法是在显微载玻片上用一种塑料膜覆盖住组织切片，用激光束活化并吸附到切割区域，组织切片部分热溶于胶片上，并能从载玻片上剥离掉残余组织。E. Von Eggeling 等将显微切割技术与蛋白质芯片结合用于肿瘤研究，比较正常组织与肿瘤组织的蛋白质表达谱差异。首先用显微切割术获取纯的肿瘤与正常组织，提取组织中的蛋白，与蛋白质芯片反应，洗脱掉未结合上的蛋白质，利用配套的分析系统检测芯片，作出蛋白质表达谱。显微切割技术能够提供来源同一的组织，非常适合于样本量要求不多的芯片分析与质谱分析。

3. 样品的标记

蛋白质芯片的实验结果除了以质谱技术为基础的直接检测法不需标记外，对样品中的蛋白质进行标记主要是为以后采用CCD照相技术或激光扫描系统等检测杂交信号时提供一个可供比较的标志物。主要的标记物是荧光物质或同位素。

（三）蛋白质芯片的杂交反应和结果检测

在蛋白质芯片上发生的主要反应类型是抗原-抗体反应，多采用酶联免疫吸附测定法（enzyme-linked immunosorbent assay，ELISA），故有人认为目前的蛋白芯片实际上就是一种高密度的ELISA反应。反应之后，再经反复洗脱，只有那些与芯片上的抗体发生反应的蛋白质能牢固粘在芯片表面。所使用的反应试剂也与常规ELISA反应的相同，包括pH 7.4的磷酸盐缓冲液（PBS，0.008mol/L，100ml蒸馏水中含 $K_2HPO_4 \cdot 3H_2O$ 182.35mg、$NaH_2PO_4 \cdot 2H_2O$ 31.2mg、NaCl 877mg）以固定样品；PBST（含0.05%吐温-20的PBS洗涤液）；含10%脱脂奶粉的PBS封阻液，含0.1% H_2O_2、0.6mg/ml DAB的PBS显色液（或其他，视不同的标记酶而异）等。实际检测过程也比较简单，在膜上进行的反应，用肉眼就可判断出结果。除了用酶标来检测外，同样也可用荧光来标记，可用荧光显微镜直接观察结果，也可以通过CCD摄像或扫描，然后用计算机来进行信息处理、结果分析。

A. Lueking等所使用的在PVDF膜上构建的活性功能蛋白质芯片上，用普通的酶联免疫吸附法（ELISA）检测，并采用了一种能与含His-tag表位的单抗作为一抗来检测PVDF膜上排列的重组融合蛋白质。非特异性结合的抗体被洗掉，滤膜与荧光标记的二抗以及底物一起孵育。紫外线激发滤膜上形成的蛋白复合体的荧光，阳性反应点与阴性反应点呈现不同强度的荧光信号，经过CCD系统扫描分析后，得到的图像再由AIDA软件包处理。该法的检测灵敏度可以达到10pg。一般，蛋白质芯片的检测浓度下限是10fmol/μl。实验还发现以错误阅读框表达的蛋白质的假阳性克隆比例低（11%）。这种蛋白质芯片确实能够用于蛋白质表达和抗体筛选，并提供了一种高通量的配件-受体反应研究手段。该芯片制作上的创新使得在利用DNA芯片进行基因表达谱分析的同时，还可同步利用蛋白质芯片进行蛋白质表达谱的分析。由于上述的cDNA文库中存在重复，如果能加以改进构建一类无重复的单克隆（uniclone），或单基因-单蛋白质系列（unigene-uniprotein set），使基因-蛋白质只在相应的阵列中出现一次，这样就可以进一步减少样品量（比如100μl血清）进行筛选。

但是，为了提高检测灵敏度，一些新的蛋白质测定所使用的技术也被采用，主要有：①质谱（MS）分析的基础上发展起来的新技术，常用且比较成熟的有基质辅助的激光解吸附离子化飞行时间质谱（matrix-assisted laser desorption ionization-time of fight-MS，MALDI-TOF-MS），以及在它的基础上改进后表面加强的激光解吸附离子化飞行时间（SELDI-TOF-MS）质谱，其原理是利用激光脉冲辐射使芯池中的待分析样品解离形成荷电离子，由于具有不同质荷比，这些离子在仪器场中飞行的时间长短不一，由此绘制出一张质谱图，该图经计算机软件处理得到模拟谱图，同时直接显示样品中各种蛋白质的分子量、含量等信息，若将它与正常人或某种疾病患者的谱图，甚至数据库中的谱图进行对照，还能够最终发现和捕获新的特异性相关蛋白及其特征。整个测定过程一般可以几十分钟内就全部完成，且方法敏感、特异性高，同时不会破坏所测定的蛋白质。②原子力显微镜（atomic force microscopy，AFM）对测定物进行表面扫描。③激光捕获显微解剖技术（laser capture microdissection，LCM）。④荧光共振能量转换。上述技术均是用来了解分析物中目标蛋白的结合变化情况。

目前的发展趋势是将各种微泵、微阀、微流路、微电泳技术运用到芯片上，以提高分析检测的速度和灵敏度。

三、蛋白质芯片的研究进展

（一）方法学研究和新技术的应用

在蛋白质芯片研究的技术领域还有许多其他方面的新进展，包括通用蛋白质阵列系统（UPA）的建立。在此系统中，使用精确、快速的 X-Y-Z 三维自动点样仪器上的 36 个毛细管为基础的印刷头点样，形成含 144 元素（4×36）的阵列。标准 ELISA 技术和扫描 CCD 检测器被用于每孔阵列中抗原的成像。

在蛋白质微阵列的形成方法上有很多的报道，如使用硅的光刻印刷或者金单层，结合使用微球传感器的微孔或喷墨到聚苯乙烯膜上。上述模式均牵涉到通过定型单个蛋白质（如 BSA、亲和素和单克隆抗体）而进行的微型免疫分析模式。

近年来，我国研究者顾宁利用分子组装技术，以琼脂糖凝胶为基质，于玻片表面制成干膜，经过碘酸钠氧化后再与戊二醛分子偶联，而成为具有双功能分子层的薄膜基片。基片表面衍生的醛基可与各类抗原（如重组多肽、磷脂和 DNA）、辣根过氧化物酶及抗体分子中的氨基形成 Schiff 碱共价连接，从而制备出各种生物分子的微阵列芯片，并优选出各种生物分子的最佳点样浓度为 0.1μg/μl 及点样量（以 50nl 为宜）。通过反复实验，成功地在芯片表面建立了酶免疫检测体系，以酶-底物（HRP-DAB-H_2O_2）作为信号显示系统，可敏感特异地检出抗原与抗体分子及酶与底物之间的相互作用。该法样本和试剂用量少，可对多种自身抗体实施高通量的平行分析，结果较为直观，具有较大的实用性和广阔的应用前景。

蛋白质芯片-飞行质谱系统的联合应用对识别特定基因克隆的蛋白质表达物、进行蛋白质水平的药物筛选、揭示蛋白激酶的作用、测定血清中的小分子物质含量等方面均比现有技术更准确、更迅速。一类新型的用配体包被的、用以捕获和分析特异性标记的蛋白质的芯片已经问世，包括采用表面增强激光解吸附离子化-飞行时间质谱技术的 SELDI 蛋白质芯片（Ciphergen Biosystems，USA）、XNA on Gold 芯片（Interactiva，Germany）、多种 BIA-CORE 芯片（Biacore，Sweden）。SELDI 和 BIA-MS 技术将亲和捕获和高分辨率质谱分析与新的、有效定性蛋白质和相互作用的配体的芯片技术结合起来，以进行表位作图、受体-配体、蛋白质-药物相互作用研究的研究。

根据芯片表面的不同化学成分，可将 SELDI 蛋白质芯片分为化学表面芯片和生物表面芯片。化学表面芯片用于检测未知蛋白，并获取指纹图谱，生物表面可显示抗原或配体的不同分子量亚型。SELDI 系统能精确检测分析物的质量，分子量范围为 0～500000。

美国 Ciphergen Biosystems 公司采用 SELDI-TOF-MS 方法，即表面加强激光解吸附电离飞行时间质谱。其原理是利用激光脉冲辐射使芯池中的分析物（即靶蛋白等）解吸附形成荷电离子，根据不同质荷比，这些离子在仪器场中飞行的时间长短不一，以分析蛋白质的分子量和相对含量，并由此绘制出一张质谱图来。该图经计算机软件处理还可形成模拟谱图，同时直接显示样品中各种蛋白质的分子量、含量等信息。若将它与常人或某种疾病病人的谱图，甚至基因库中的谱图进行对照，还能够从中发现和捕获新的疾病特异性相关蛋白及其特征。整个测定过程一般可以在几分钟内就全部完成，十分迅速且方法敏感，特异性高，同时不会破坏所测定的蛋白质。该公司通过该项技术，成功地检

测到了纳摩尔的淀粉样β蛋白（Aβ），而且定量地评价了Aβ40与Aβ42的比例，为进一步研究老年性痴呆的发病机理、发展生物标志物及γ-1-42分泌酶的特异抑制剂提供了基础。

（二）抗体芯片

免疫传感芯片已经研制成功，从而可以同步检测临床分析物。虽然目前还是低密度模式（6×6模式），但发展为高通量的潜能极大。在这类芯片上，捕获抗体和分析物以阵列的形式被加样到有十字交叉样流动室的显微玻片上。也有文献报道可以利用修饰的氨基接受表面以微模水凝胶“印章”的方式制备芯片，将蛋白质固定为亚单层，最重要的是芯片中抗体是有活性的。实验结果可以用^{125}I标记或原子显微镜显示，也可以通过荧光标记和CCD为基础的光学读数装置来完成。这类芯片由于涉及自动化成像分析和微流体力学技术，除了抗体分析外，还可能用于酶活力或其他分析的一种方法学模式。

大量的重组蛋白质也被用来制作抗体芯片，如用噬菌体表面展示的抗体库。虽然目前这种高通量生成的特异性抗体比较昂贵并需要大量人力，但是其创意却十分新颖和具有吸引力。

第四节　生物芯片在医学中的应用

一、基因芯片在医学中的应用

（一）微缩实验室芯片

生物分析仪器及过程微型化和集成化是20世纪90年代生命科学、化学和微电子学等领域研究和开发的重要方向之一。分析仪器的体积将从置于地板上→安放在桌面上→便携式掌上操作。以DNA测序为例，常规的分析过程包括样品处理、DNA提取、酶切、PCR扩增和电泳分离等步骤。如果将上述分析功能集成和浓缩到只有火柴盒大小的芯片上进行，在玻璃或硅片上光刻印刷制作各种微阀、加热器、温度传感器、微流路、反应室等，一张芯片就如一个综合实验室（lab on chip），将极大地加速DNA分析的进程。目前，已制备的微流路分析仪的“雏形”芯片有：纳米级DNA分析芯片；PCR芯片及PCR-毛细管电泳芯片和细胞过滤-PCR芯片等，可在体积为4.5μl的硅玻璃芯片上实现从样品中分离白细胞以及从分离的白细胞中直接进行PCR扩增和电泳分析的功能；蛋白质分析芯片，在这种芯片上能实现血液样品与检测试剂的混合、反应、分离和分析等一系列功能；样品处理芯片（Affymetrix公司），这种芯片尺寸稍小于信用卡，可插入到一个工作站内使用，并执行7种不同的程序从血液中破碎细胞、提取DNA、逆转录、荧光标记、裂解RNA、样品扩增、处理、稀释等功能，最后将处理好的样品溶液转送至DNA芯片进行杂交。该芯片还备有微型加热器，以保证反应池中各种热循环及冷循环反应温度的需要；电寻址导向式芯片（Nanogen公司），采用介电电泳的原理可将大肠杆菌从人体血细胞的混合物中分离出来，并用蛋白酶裂解处理，所提取的大肠杆菌DNA经电寻址导向式杂交方式在DNA芯片上完成杂交分析；

其他与生物传感器相连的各种整合型功能芯片亦在不断地问世。

（二）基因突变分析

基因诊断是 20 世纪 70 年代末期伴随着基因重组技术而发展起来的一项以分析核酸中碱基序列特征为基础的新的诊断策略。基因诊断的核心技术是 DNA 分子杂交和 PCR 技术。生物芯片的出现使突变检查变得更加规模化和程序化。

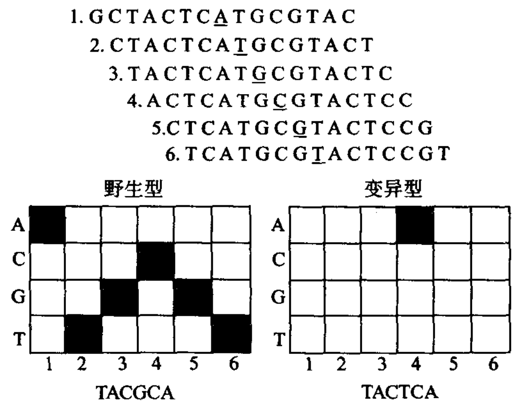


图 13-9 4L 砌砖方法用 ODN 芯片作 DNA 再测序以分析鉴定点突变

用生物芯片检测基因突变的一个成功的例子就是用 4L 砌砖方法检测人线粒体 DNA 的变异。首先按照图 13-9 合成 15mer 的寡核苷酸（ODN）探针。

芯片分为 1～6 列，每列均原位合成或点加序列为 1～6 的 15mer 寡核苷酸探针。在每一列中探针序列只有中间 1 个有下划线的核苷酸不同，当这样的芯片与带荧光的样品杂交时，可以在芯片中出现序列特异的杂交结果（黑色方块）。荧光测序结果很明显地显示出野生型（序列 TGCGCA）及变异型（序列 TACTCA）。只要有 1 个碱基的突变，引起的错配可以使与点突变相邻的变异序列杂交信号明显减弱。

将人线粒体 DNA 的全序列分为 16000 个 20mer 的寡核苷酸于芯片上，所形成的 64000 个点阵的芯片可用于检测整个线粒体基因组中的变异，只需作 1 次杂交。曾在 180 个变异体中用 DNA 芯片成功地检出其中 179 个变异体。

目前，DNA 芯片技术已成功地用于人酪氨酸激酶基因、β-珠蛋白基因、HIV-1 蛋白酶和逆转录酶基因、囊性纤维变性基因、*C-Ras*、*C-Myc*、*C-Erb B2*、*p53*、*Rb-1*、*p15*、*p16* 等多种癌基因或抑癌基因突变的检测。理论上各种遗传病及肿瘤相关基因的变异，如血友病、杜氏肌营养不良症、地中海贫血、异常血红蛋白病、苯丙酮尿症的基因突变，都可以用相似策略的芯片鉴别出来。

（三）克隆选择及文库筛选

生物芯片与 DNA 文库和噬菌体文库结合，进行文库的分类和筛选。

（1）cDNA 文库分类　将人体特定组织或器官的 cDNA 文库中的克隆扩增后，以微点阵方式固定在杂交膜上，选用特定的探针系统与之杂交，再以计算机软件分析，就可快速、低廉、高效地对一个大容量 cDNA 文库中代表不同基因的 cDNA 克隆进行分类和分析，让人们更清楚简便地了解组织或器官在发生、发展、老化及癌变过程中基因表达水平，以及整个人体内基因网络的变化。

（2）噬菌体文库分析　BIACORE 技术属于生物芯片中的一种，它依据表面等离子体共振原理，分析生物分子间的相互作用。将其与噬菌体抗体文库相结合，将目标抗原固定在生物芯片上，加所筛选的抗体文库到缓冲液中，以液体流动形式不断地与固定化抗原作用。通过改变洗脱液的流速、收集不同时间点流出的噬菌体，就可得到不同亲和力的小肽分子，洗脱时间越长得到的分子亲和力就越高。通过计算已知的抗原与抗体相互作用的动力学常数，可人为调控二者作用的时间，最终筛选到高亲和力的小肽。

（四）测序芯片

基于杂交原理的DNA芯片测序法（sequencing by hybridization，SBH）可分为两类：①一张芯片上固定众多的寡核苷酸点阵，它们的碱基序列不同、变性温度（T_m）不同、最佳杂交条件不同，在同一张芯片上固定后进行杂交，很难找到适合于所有的ODN的统一温度，故易出现假阳性和阴性；②使用5mer、6mer、7mer等较短的ODN时，有利于错配分辨，但稳定性较差，易引起杂交信号的下降，特别是末端碱基的错配，尤以T-G、G-A错配较难检测。

测序芯片可应用于基因图谱的绘制、DNA测序、基因多态性分析等方面。现以DNA序列测定为例作一介绍。如果寡核苷酸的长度为N个碱基，在芯片上可建立4*n*个长度为N序列的阵列集合，而待分析的靶序列的长度大约为点阵上寡核苷酸数目的平方根。例如一个含8个ODN的65536个点阵的芯片，可测定约200个核苷酸的序列；而含1×10^6个12个寡核苷酸的点阵可用于测定约1000个碱基的序列。用这类芯片对已知序列的含922bp的*Eco*RⅠ/*Bgl*Ⅱ人基因片段（含*β*-干扰素基因）作验证性分析，测定了其中从627～726bp的片段。使用72个8mer和21个9mer的ODN，杂交结果表明，完全匹配的DNA杂交信号比错配的DNA杂交信号强得多，而且分析结果完全相符。

但是该方法对很大的未知DNA片段进行测定，需要较长的ODN点阵，容量相当大，现行方法还难以做到，但仍有发展前途。

（五）基因多态性分析

多态性分析最广泛的方法是用ASO探针（等位基因特异的寡核苷酸探针）进行差示杂交。现以直接荧光检测方法（direct allele specific fluorescence targeting，DAFT）说明之。按下列步骤进行：①在玻璃片上制备ASO平面点阵；②合成和纯化荧光标记的单链PCR探针；③标记的PCR产物与ASO点阵杂交；④荧光扫描点阵杂交结果，鉴定基因的多态性。用DAFT检测方法测得了人类酪氨酸酶基因第4个外显子的5个点突变位点。

基因组单核苷酸多态性分析（SNP）由于其包含的信息量极大，加上平均每1000个碱基就有一个单核苷酸多态性变异，因此近年来基于SNP检测的基因芯片受到重视。Chee等应用包含135000个25mer寡核苷酸探针的微阵列基因芯片，检测出人类线粒体基因组（16.6kb）的SNP位点。结果从10个检测样品中找到了505个SNP标志位点。近年也报道了将凝胶测序法与高密度基因芯片技术结合，对2.3Mb的人类基因组长度进行了SNP筛选，最终确定3241个SNP位点，其中2227个位点被用来作为构建基因图的路标。

（六）基因表达谱分析

对来源于不同个体、不同组织、不同细胞周期、不同发育阶段、不同病变、不同刺激（包括不同诱导、不同治疗阶段）下的细胞内的mRNA或逆转录后产生cDNA与表达谱基因芯片进行杂交，可以对这些基因表达的个体特异性、组织特异性、发育阶段特异性、分化阶段特异性、病变特异性、刺激特异性进行综合的分析和判断，迅速将某个基因或几个基因与疾病联系起来，极大地加快这些基因功能的确定，同时进一步研究基因与基因之间的相互作用。

目前，国内已有多家生物高科技公司制备了各类cDNA表达谱芯片，诸如癌基因和抗癌基因芯片，信号转导芯片，细胞周期检测芯片，细胞凋亡检测芯片，细胞因子、生长因子及生长因子受体检测芯片等，为基因表达的整体研究提供方便。

（七）微生物菌种鉴定、致病机理及耐药性分析

DNA 芯片可根据细菌或病毒基因组分析结果进行基因分型、菌种鉴定及流行病学研究；还可以对微生物的重要基因，如毒力基因、耐药基因、致病因子等进行监测，寻找致病机制及耐药机理，为临床治疗提供依据。

从基因芯片的研究中发现 90%～95%的结核杆菌对利福平的耐药性与 RNA 聚合酶的 β 亚基（*RpoB* 基因）有关，3534bp 编码区中存在有 81bp 的突变区域。用基因芯片技术对 44 株临床上对利福平有抗性的结核杆菌进行 *RpoB* 基因的突变分析，涉及到其中 41 株 8 个密码子 12 种突变，包括错义突变、缺失突变等。还对 2 个以上不同的耐药性基因（如 *KasA*、*PncA*、*Emb* 等）同时进行基因芯片探针杂交测定。这种结核杆菌多个耐药性基因的检测芯片现在已有商品出售。

（八）药物筛选芯片

利用基因芯片可以比较不同组织细胞中基因的表达模式、研究正常组织与病理组织基因表达差异、建立模式生物细胞中的基因表达模型、建立病原体基因的表达模型、研究药物处理细胞后基因表达变化等。由此可见，基因芯片对于药物靶标的发现、多靶位同步高通量药物筛选、药物作用的分子机制、药物活性及毒性评价方面都有其他方法无可比拟的优越性，能够从基因水平解释药物的作用机制，可以用基因芯片分析用药前后机体的不同组织、器官基因表达的差异，国外几乎所有的主要制药公司都不同程度地采用了基因芯片技术来寻找药物靶标，检查药物的毒性或副作用。

药物基因组学（pharmacogenomics）是在基因组学的基础上研究不同个体对药物反应的差异以便对不同的基因型患者实行个体化给药，从而将药物的药效充分发挥而使不良反应减少到最小。将 DNA 芯片技术应用于药物基因组学，一方面可以加速药物基因组学的发展，主要是利用 DNA 芯片进行基因功能及其多态性的研究，以确认与药物效应及药物吸收、代谢、排泄等相关的基因，并查明这些基因的多态性；另一方面 DNA 芯片利用药物基因组学的研究成果，根据基因型将人分群，以实现药物基因组学研究的目的和价值。

中药在现代药物研究中日益受到重视，但中药种类多，其药效受产地、气候、收获季节等多种因素的影响，并且还存在假冒伪劣中药材的情况。生物芯片在名贵名药材鉴定、药效分析、药物有效成分分析等方面也有极大的应用潜力。

二、蛋白质芯片在医学中的应用

（一）蛋白质组-基因组链接

使用单蛋白质系统（uniprotein set）中的蛋白质所建立起来的蛋白质芯片将在基因组学和蛋白质组学方法之间架立了桥梁。基因组学方法是利用高密度 cDNA 芯片绘制一个细胞中 mRNA 库的谱图，能作为基因表达的第一层次直接反映了基因活动、基因表达状况并借此发现新的疾病相关的基因；蛋白质组学方法则涉及到质谱鉴定后肽谱、序列标签或者由个别蛋白质裂解产物片段-离子指纹图的建立，并通过与序列数据库的比对鉴定这些个别蛋白质。这两个方法之间的链接可以通过使用单克隆系列中的蛋白质来实现。这意味着，差异表达的基因通过 cDNA 微阵列被识别后，同一 cDNA 克隆还可以在不同细胞系统或者通过体外转录和/或翻译进行同步蛋白质表达分析，这些被表达的蛋白质被制成与 cDNA 克隆相对应的蛋白质芯片微阵列，然后用抗体、核酸、小分子配体或标记药物与芯片上的克隆表达产物反应进行高通量的筛选；甚至可以使用质谱测定来鉴定阳性的反应

物。因此，可以说蛋白质芯片促进了单个克隆的DNA系列信息与对应蛋白质产物的之间联系，使得翻译的基因产物直接经得起高通量实验的检验，并建立起蛋白质表达和DNA系列数据之间的直接联系。在此基础上，再回归到整体基因组水平的研究，以达到对生命活动本质的深刻认识。

为了做到这一点，P. Uetz等利用酵母细胞构建了第一个全基因组的活性蛋白质芯片以及A. Lueking等首先将人胚脑cDNA表达文库克隆至载体上，诱导原位表达包含His-tag的融合蛋白质，然后将微量滴定板中的表达蛋白点在PVDF膜上，建立蛋白质芯片。当然，要完善此类蛋白质芯片还涉及到单基因系列（unigene sets）、微阵列和单蛋白质系列（uni-protein sets）以及质谱分析中的“最少蛋白质识别物”（minimal protein identifiers，MPIs）等概念，在此不做多述。

（二）高通量的蛋白质表达筛选及药物研究

W. K. Buss等用类似的技术完成了高通量的蛋白质表达筛选。首先，将cDNA文库克隆到原核表达载体，并用仪器将其克隆菌落排列在微量滴定板上，再转移到PVDF膜。用抗体与膜上的蛋白质反应，筛选反应阳性的克隆，作为第一轮筛选。第二轮筛选：小规模的蛋白质表达（指上轮的阳性克隆）在微量滴定板上进行，再用SDS-PAGE、亲和层析、质谱分析表达产物的大小、产量、纯度、溶解度。对克隆的表达水平同时进行分析，以期发现最适于高通量结构分析与功能筛选的克隆。发现66％的克隆包含插入片段，并且插入方向正确，其中又有64％的克隆包含了完整的插入片段，能够编码一个完整的蛋白质。第三轮筛选：在微量滴定板上对蛋白质功能进行分析，测定了96个克隆的细菌裂解产物表达GADPH的活性。通过三轮筛选，作者建立了一个人类蛋白质表达克隆目录，为今后的结构与功能分析提供了基础。

高通量的蛋白质表达筛选最实际的应用就是寻找新的药物，通过用药前后对正常组织或细胞及病变组织或细胞中大量相关蛋白表达的变化进行比较，充分了解细胞信号转导和代谢途径，进而发现一组疾病相关蛋白作为药物筛选靶标、发现新的药物、了解药物体内代谢状况和毒性（可达10pg级）并进行安全性的评价。另外，蛋白质芯片技术不仅可以研究各种化合物与其相关蛋白质的相互作用，还可以在对化合物作用机制不了解的情况下，直接研究疾病的蛋白质表达谱，从而将化合物的作用机制与疾病联系起来，并进一步建立外源化合物与疾病蛋白质表达谱库，为新药开发和各种药理研究提供大量数据。

（三）蛋白质和蛋白质之间相互作用的研究

蛋白质芯片通过标记可能存在反应性的蛋白质（如配体、酶等），使用特异性抗体或质谱测定，检测和筛选多种蛋白质之间的相互作用。

1. 抗原-抗体相互作用

用于高通量的抗原-抗体反应筛选的抗体微阵列，首先将细菌文库的单菌落排列在第一张膜上，膜上的细菌菌落可以分泌蛋白产物——抗体。第二张膜上印有抗原，能够与第一张膜上的抗体结合。将第一张膜压在第二张膜上，抗体就被第二张膜上的抗原黏附，于是第二张膜就具有了第一张膜上的所有抗体阵列。利用此抗体阵列，再与细胞裂解产物进行反应，筛选存在于细胞裂解物中的所有能与抗体阵列反应的抗原，研究细胞内的各种蛋白质表达水平。这种抗体芯片的突出优点是抗体阵列能够与细胞裂解物直接反应，比目前通用的分离复杂成分中蛋白质组的二维电泳要方便和快捷得多，而且得到的信息也更多，只要进一步提高灵敏度，很有发展前途。

2. 酶-底物相互作用

Heng Zhu 和 Michael Snyder 报道了蛋白质微孔芯片用于激酶底物磷酸化的研究。将微孔芯片的表面用交联剂 GPTS（3-glycidoxypropyltrimenthoxysilane）活化并在 130℃保温 1h，再将蛋白质底物加入微孔中，放在湿度盒中并置于冰上平衡 3h。微孔用 1%的 BSA 封闭后，用 TBS（tris-buffered saline）洗涤，然后加入激酶、^{33}P-γ-ATP 和缓冲液，接下来在 30℃恒温 30min，洗净芯片，在 MDP（molecular dynamics phosphoimager）中自显影成像，见图 13-10(a)。图 13-10(b) 是用 poly(Glu-Tyr) 作底物得到的磷酸化结果放大图。

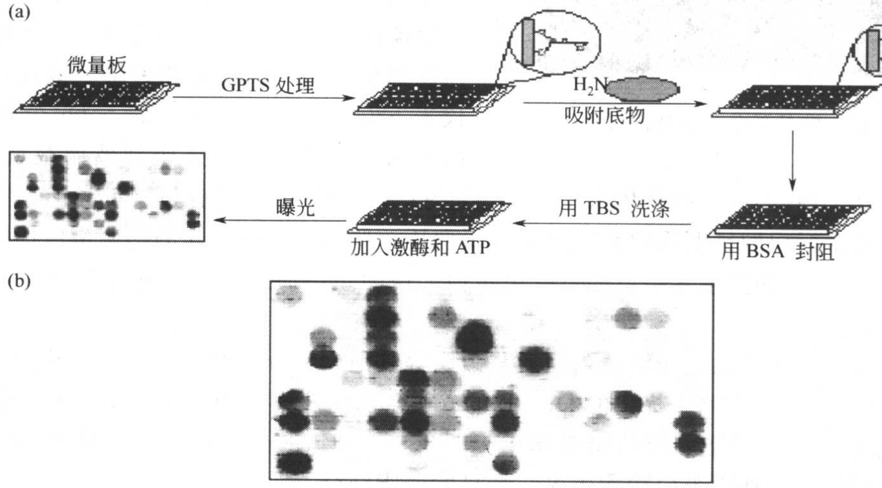


图 13-10　蛋白质芯片对激酶-底物作用的研究

3. 蛋白质-配体相互作用

这是蛋白质芯片最常用的研究目标。Thomas kodadek 报道了用蛋白质功能芯片探测蛋白质-药物之间的相互作用。分别将 Protein G、p50 和 FKBP 上的 FRB（FKBP-rapamycin binding）结构域点在化学处理后的玻片上，每个靶点一式四份。然后，用不同的荧光标记蛋白进行探测。如图 13-11 所示，A 板加入能与蛋白 G 结合的 Bodipy-FL-IgG 进行探测；B 板用能与 p50 结合的 Cy3-labeled-IkBa 探测；C 板用雷帕霉素及能与 FRB 结合的 Cy5-labeled-FKBP12 探测；D 板除了不加雷帕霉素外，其他与 C 板一样；E 板是在雷帕霉素存在条件下，同时用三种标记蛋白探测，实验获得非常满意的结果。

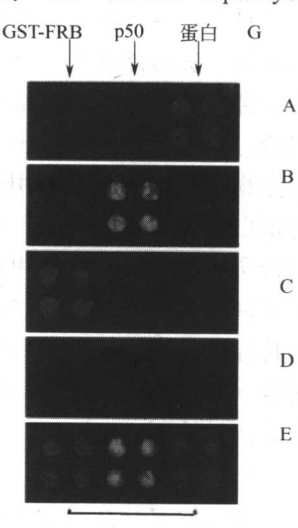


图 13-11　蛋白芯片探测蛋白质之间的相互作用

现在已经有了很先进的用于捕获和分析特异性标记蛋白质配体的芯片，包括 SELDI（表面加强的激光解析离子化技术）蛋白质芯片和多种 BIACORE 芯片。SELDI 和 BIA-MS 技术将亲和捕获和高分辨率质谱分析与新的、有效定性蛋白质和相互作用的配体的手段结合了起来，可应用于诸如表位作图、受体-配体或蛋白质-药物相互作用的研究。

4. 蛋白质翻译后修饰的检测

例如，许多细胞功能通过可逆磷酸化进行调节，特定磷酸化位点及磷酸化程度的鉴定对于理解这些修饰如何调节受体活

性至关重要。蛋白芯片与MS技术结合可快速、灵敏地用于翻译后修饰的测定。

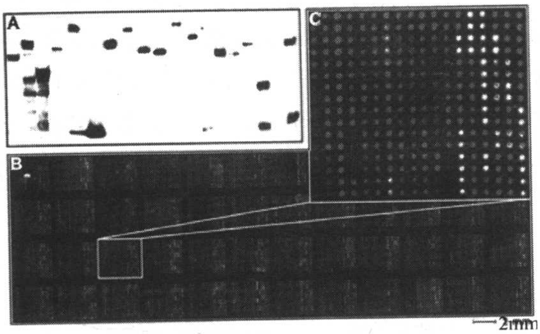


图 13-12　酵母蛋白芯片分析信号

耶鲁大学的Micheal Snyder等构建了一组酵母蛋白质组芯片来筛选一些蛋白质修饰后的生化活性。他们将5800个酵母开放阅读框克隆到酵母高拷贝表达载体上，使用可诱导的GAL1启动子，在酵母中表达。在96孔板中，利用谷胱甘肽琼脂糖珠从酵母提取物中迅速纯化出1152个蛋白质样品。利用免疫印迹方法随机抽样其中的60个样品来监测纯化蛋白质的质量和数量。最终，得到代表5800种酵母开读框的6566个蛋白样品。用点样仪将它们点印到玻璃载玻片上，制成蛋白质芯片。用抗GST的抗体来检测，超过93.5%的蛋白样品有显著信号。在钙存在的条件下用生物素标记的钙调蛋白来检测酵母蛋白质，鉴定出6种已知的钙调蛋白标靶，以及1种结合Cy-3标记的抗生素蛋白的链霉菌蛋白，说明蛋白质芯片有可能鉴定出许多转录后修饰蛋白。

酵母蛋白质芯片还可以用于筛选蛋白质与磷脂等其他生物分子的相互作用。利用含有磷脂酰胆碱的脂质体、1%生物素标记的DHPE的5类磷脂酰肌醇（PI）脂质体和1类缺乏PI的脂质体来检测蛋白质-磷脂之间的相互作用。图13-12显示的杂交结果表明体内许多膜蛋白包括完整的膜蛋白和外周蛋白均能优先与特异的磷脂结合。

（四）在疾病诊断方面的应用

除了上面已介绍的用PVDF膜蛋白质芯片对各类病毒性肝炎进行诊断外，蛋白质芯片在临床方面一个重要的应用即是对肿瘤的早期诊断。另外，还有HCV-TP-HIV三联蛋白芯片，由于酶标记灵敏度太低，多数采用荧光标记或生物素亲和素标记的方法。

虽然，各种肿瘤标志物在临床上已应用多年，但在应用过程中，单指标检测存在着特异性不强、灵敏度低等不足等问题。临床上常将几项相关的标志物组成联合标志物组，同时对某一肿瘤进行检测，以提高肿瘤诊断的准确性。在血清肿瘤标志物的测定方法多以酶联免疫实验、放射免疫实验、电化学发光等为主，但这些方法同样存在血清量较大、只能分析单项标志物等缺点。如果利用蛋白质芯片的高通量、并行化、微量化的特点，发展“多肿瘤标志物蛋白质芯片诊断系统”无疑是蛋白质芯片走向临床应用的正确方向。上海数康生物科技有限公司研制开发的含有甲胎蛋白（AFP）、癌胚抗原（CEA）、糖原199（CA199）、糖原242（CA242）、糖原125（CA125）、糖原153（CA153）、神经元特异性烯醇化酶（NSE）、前列腺特异抗原（PSA）、游离前列腺特异抗原（f-PSA）、铁蛋白（FER）、β-人绒毛膜促性腺激素（β-HCG）、人生长激素（HGH）12个肿瘤标志物的蛋白质芯片很好地用于肿瘤病人的血清学诊断。

（五）展望

蛋白质芯片在医学中的应用是绘制体液中DNA、蛋白质、肽和抗体的图谱，它被认为能够促进对健康和疾病的理解。例如，通过筛选血清、脑脊液和其他正常和疾病的体液可能识别与疾病相关的蛋白质。蛋白质芯片还可以描绘和监测疾病状态下，诸如自身免疫性疾病的体内蛋白质表达状况，通过监测特殊治疗的效应或者抗体对疫苗的反应，为医生提供与诊

断或预后相关的抗原信息。也可能通过这样的筛选去发现一些新的靶标。所以，蛋白质芯片在医学领域有着广泛的用途。

蛋白质芯片技术从产生至今已有了很大的发展，但与基因芯片相比较还处在起步阶段，无论在芯片的制备、具体应用过程以及结果的检测方面还有很多的不足。因此，今后在减低成本、简化芯片的制作工艺（如探针制备和固定、样品标记等）、增加稳定性和可重复性、发展便携式的信号检测仪器等方面有所突破，芯片的应用才会走向普及。相信随着人们对蛋白质结构和功能认识的不断深入，以及其他辅助学科和技术的发展和成熟，蛋白质芯片技术会在生命科学领域发挥越来越重要的作用。未来会制作出敏感性和特异性更强的固定有多种大量活性蛋白的新型蛋白质芯片。

（杨瑜珍）

第十四章　生物信息学技术

生物信息学是综合运用生物学、数学、物理学、信息科学以及计算机科学等诸多学科的理论方法的崭新交叉学科，它的研究对象是生物学的数据，而它所采用的方法，则是从各种计算技术衍生而来。广义地说，生物信息学从事生物信息的获取、加工、储存、分配、分析和释读，并综合运用数学、计算机科学和生物学工具，以达到理解数据中的生物学含义的目的，并揭示“基因组信息结构的复杂性及遗传语言的根本规律”，解释生命的遗传语言。它已成为整个生命科学发展的重要组成部分，成为生命科学研究的前沿。

第一节　概　　述

生物信息学的内涵丰富，其核心内容至少包括基因组信息学、蛋白质组信息学和代谢调控信息学三大部分。基因组信息学指基因组信息的获取、处理、存储、分配和解释，目的是确定全部基因在染色体上的确切位置以及各 DNA 片段的功能；蛋白质组信息学对细胞和组织中的蛋白质结构、表达、定位以及各蛋白质之间的相互作用信息进行处理和解释，最终确定各蛋白质的功能及相互作用关系，蛋白质组学研究不仅能为解释生命活动规律提供基础，还能为多种疾病机理的阐明提供理论根据和解决途径；代谢调控信息学的目的是阐明基因表达的调控机理，它通过研究生物分子在基因调控中的作用，来了解人类疾病诊断、治疗的内在规律。

一、生物信息学的诞生和发展

生物信息学的萌生可追溯到计算机初创期的 1956 年，当时在美国田纳西州的 Gatlinburg 召开了首次“生物学中的信息理论讨论会”。然而，作为一门正式的学科却只有 20 多年的历史，因此，它是一门相当年轻的学科。在沉寂了多年之后，伴随着 20 世纪八九十年代计算机技术的迅猛发展，它才得以迅速发展。

随着人类基因组计划的全面展开和深入，相关信息出现了爆炸性增长。自 1995 年科学家破译全长为 180 万个核苷酸的流感嗜血杆菌基因组以来，至 2004 年 2 月已有 170 多个模式生物如酵母、线虫、果蝇、拟南芥等的基因组被测定。根据国际数据库的统计，1990 年核酸碱基数目为 4000 万，1995 年核酸碱基数目为 2.5 亿个，2000 年核酸碱基数目为 58 亿个，到 2004 年 2 月这一数目已超过 200 亿。要对这些海量数据进行处理分析就必须发展新的分析理论、方法、技术和工具，也必须依赖计算机的信息处理，这正是生物信息学迅速发展的基础。

二、生物信息学研究的基本方法

生物信息学研究方法包括生物学数据库的建立、生物学数据的检索（收集和筛选）、处

理（编辑、整理、管理和显示）及利用（计算、模拟）。

1. 生物学数据库的建立

生物学数据库的建立通常是由专门的机构来完成。这些机构包括一些国家支持的非营利性机构和一些知名大学的研究院，另外有部分营利性机构开设了一些商业性数据库（如MDL），一些大的制药公司和生物领域相关公司也都建立了自己的数据库，当然也有某些实验室为了研究工作的需要建立适合自己的小型数据库。目前，最为著名的三大核心数据库是PDB生物大分子结构数据库、SWISS-PROT蛋白质序列数据库和GenBank核酸序列数据库。

2. 生物学数据的检索

生物学数据库的检索对象是多方面的，按照不同的需要检索不同类型的数据库。例如，查找相关文献就应该选择文献数据库（如Medline），查询DNA序列则选择核酸序列数据库（如GenBank）。生物学数据库的检索包括两个层面——收集和筛选。检索可以登录相关数据库，利用数据库本身自带的检索工具进行检索，也可以利用很多开发好的检索软件进行检索，如“Reference Manager”可对文献进行检索和管理。

3. 生物学数据的处理

生物学数据处理的对象应当分为两类数据：一类是实验中产生的数据；另一类则是在数据检索中查得的数据。对于这两类数据的处理，基本上是相同的。考虑到进行计算时的数据库格式需要，首先必须对数据进行格式编辑。对数据还必须进行分类、整理，这样有助于查询，如果数据量较大时，往往需要建立一个自己的小型数据库并进行管理。

4. 生物学数据的利用

生物学数据的利用，其核心是计算生物学中产生的大量算法。各种统计模型和算法纷纷使用到计算生物学中，如隐马尔可夫模型（hidden Markov model，HMM）、最大似然模型（maximum likelihood model，MLM）、遗传算法（genetic algorithm）、人工神经网络（artificial neural network）。只有利用这些算法才能对数据进行分析，包括核酸或蛋白质序列相似性比对分析、蛋白质刚体结构比对分析、外显子预测、分子进化，以及基因组分析等。

三、生物信息学的目的意义

分子生物学数据令人惊愕的数据积累以及对它们内部神秘而精细的作用模式的揭示形成了当前数以百计的生物信息数据库。它们各自按一定的目标收集和整理生物学实验数据，并提供相关的数据查询、数据处理的服务。随着互联网的普及，这些数据库大多可以通过网络访问或者通过网络下载。

一般而言，这些生物信息数据库可以分为一级数据库和二级数据库。一级数据库的数据直接来源于实验获得的原始数据，只经过简单的归类整理和注释；二级数据库则在一级数据库、实验数据和理论分析的基础上针对特定目标衍生而来，是对生物学知识和信息的进一步整理。一级和二级数据库之间存在着本质的差别，一级数据库记录实验结果和一些初步的解释，而更进一步分析工作的结论只能从二级数据库中查找到；因此，使用数据库时，对数据库要正确定位。国际上著名的一级核酸数据库有GenBank数据库、EMBL核酸库和DDBJ库等；一级蛋白质序列数据库有SWISS-PROT、PIR等；蛋白质结构库有PDB等。国际上二级生物学数据库非常多，它们因针对不同的研究内容和需要而各具特色，如人类基因组图谱库GDB、转录因子和结合位点库TRANSFAC、蛋白质结构家族分类库SCOP等。

第二节 常用核酸序列数据库

1. GenBank 数据库

GenBank 数据库（http：//www. ncbi. nlm. nih. gov/GenBank/）包含了几乎所有已知的核酸序列和蛋白质序列，以及与它们相关的文献著作和生物学注释。它是由美国国立生物技术信息中心 NCBI 建立和维护的。它的数据直接来源于测序工作者提交的序列、由测序中心提交的大量 EST 序列和其他测序数据以及与其他数据机构协作交换数据而来。GenBank 每天都会与欧洲分子生物学实验室（EMBL）的数据库和日本的 DNA 数据库（DDB）交换数据，使这三个数据库的数据同步。

GenBank 库里的数据来源于约 130000 个物种，其中 56％是人类的基因组序列（所有序列中的 34％是人类的 EST 序列）。每条 GenBank 数据记录包含了对序列的简要描述、它的科学命名、物种分类名称、参考文献、序列特征表以及序列本身。所有数据记录被划分在若干个文件里，如细菌类、病毒类、灵长类、啮齿类以及 EST 数据、基因组测序数据、高通量基因组序列数据等 16 类，其中 EST 数据等又被分成若干个文件。完整的 GenBank 数据库包括序列文件、索引文件以及其他有关文件。索引文件是根据数据库中作者、参考文献等字段建立的，用于数据库查询。

随着国际互联网的发展，目前检索 GenBank 的最佳方式是通过万维网进行。在 NCBI 网站上可以轻松地检索 GenBank 数据库（图 14-1，NCBI 检索界面）。

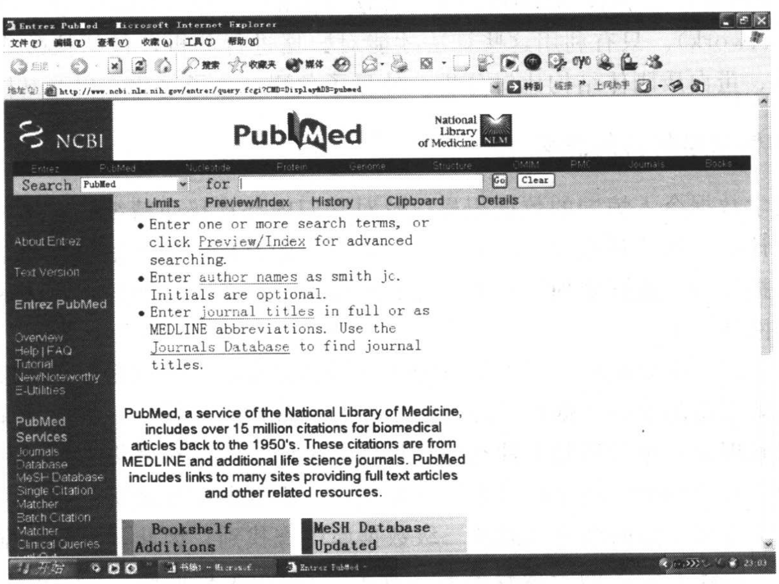


图 14-1 NCBI 检索界面

对核酸序列库的相应字段进行限定性检索，需通过点击检索窗口下方的“Limits”链接按钮，进入限定性检索界面（图 14-2）。首先在检索框中输入检索词，然后在“Search

Field”下拉菜单选项中选择相应的检索字段。其检索字段包括：accession（收录登记号）、all fields（所有字段）、author name（作者名）、EC/RN number（CAS化学物质登记号或酶编号）、feature key（基因特征词）、gene name（基因名）、issue（发表期刊期号）、journal name（发表文献的期刊名称）、key word（关键词）、modification date（修改日期）、organism（生物体）、page number（文献发表页码）、primary accession（原始登记号）、properties（特性）、protein name（蛋白质名称）、publication date（文献出版日期）、seq ID string（授予序列号的核酸序列）、sequence length（序列长度）、substance name（物质名称）、title word（题名词）、text word（文本自由词）、volume（文献发表期刊卷号）、默认为全部字段（all fields）。

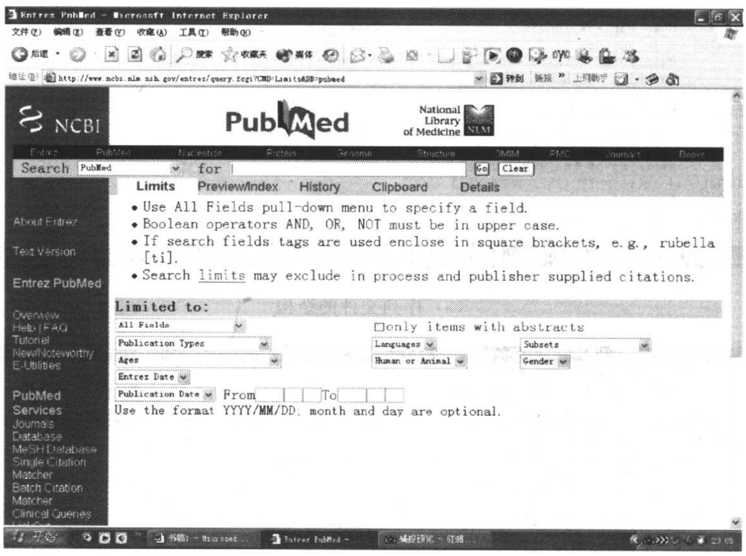
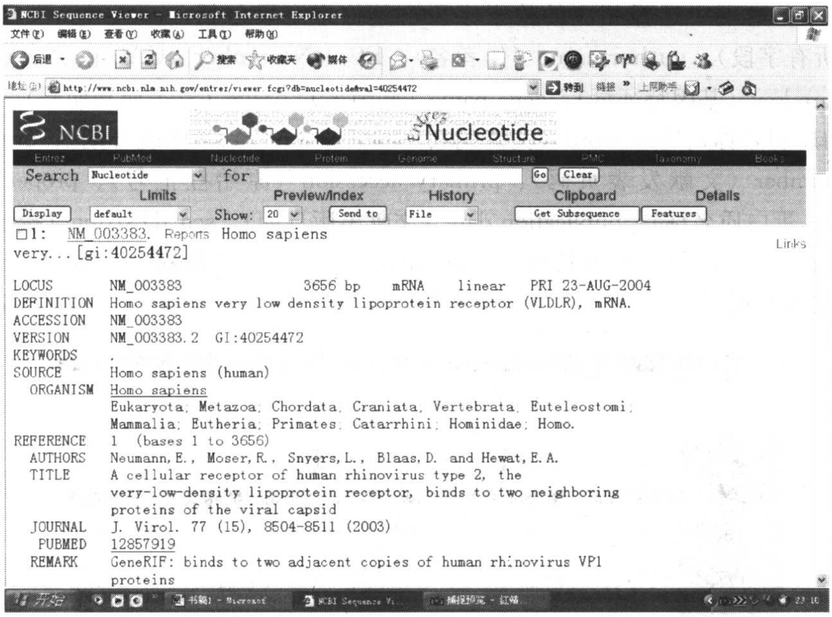


图 14-2　NCBI 限定性检索界面

GenBank中最常用的是序列文件。序列文件的基本单位是序列条目，包括核苷酸碱基排列顺序和注释两部分。目前，许多生物信息资源中心通过计算机网络提供该数据库文件。下面，我们介绍序列文件的结构（图 14-3，以 VLDLR 的序列为例）。序列文件由单个的序列条目组成。序列条目由字段组成，每个字段由关键字起始，后面为该字段的具体说明。有些字段又分若干子字段，以次关键字或特性表说明符开始。每个序列条目以双斜杠“//”作结束标记。序列条目的格式非常重要，关键字从第一列开始，次关键字从第三列开始，特性表说明符从第五列开始。每个字段可以占一行，也可以占若干行。若一行中写不下时，继续行以空格开始。

（1）LOCUS（序列名称）　是该序列条目的标记，或者说标识符，历史上曾用这个名称来表示本记录描述的基因座，提交者和数据库工作人员花费了无数的时间来设计这一名称。这一成分开始于一个英文字母，总长度不能超过 10 个字符。第二个字符以后可以是数字或字母，所有字符均要大写。LOCUS名称在以前是最为有用的，那时大多数DNA序列记录只表示一个基因座，这样在GenBank中寻找一个可以用少数几个字母和数字来代表生物体的独特的名字是很容易的事。经典的例子包括 HUMHBB（人 β-珠蛋白基因座），或 SV40 [猿猴病毒（拷贝之一，因为存在许多拷贝）]。为了可用起见，LOCUS名称在数据库中必须是唯一的。因为几乎所有有意义的命名符都被使用过了，所以今天LOCUS名称已不再是一个有用的成分，但仍有许多软件包依赖于一个独一无二的LOCUS名称，所以数据库还不



(a) 序列文件的结构



(b) 序列文件的结构

图 14-3 VLDLR 序列文件结构

能将其彻底去掉。可行的办法是代之以一个独一无二的词，最简单的是用一个保证不会重复的检索号码，以满足对 LOCUS 名称的要求。该字段还包括其他相关内容，如序列长度、类型、种属来源以及录入日期等。

(2) DEFINITION（序列简单说明） 在 GenBank 记录中用以总结记录的生物意义。这一行将出现在 NCBI 的 FASTA 文件中，这样任何人进行 BLAST 相似性搜索时都会看到这些信息。生成这一行时要非常小心，因为许多记录生成工作可以部分地自动进行。所以，数据库工作人员要检查这一行以保证信息的一致性和有效性。但是，用一行文字来说明生物背

景并不总是可行的，对此不同的数据库采用了各自的解决方法。其中有一些共识，并且每个数据库也都了解其他数据库的解决方法，并尽力相互一致。

(3) ACCESSION（序列编号）　具有唯一性和永久性，如本例中代码 NM _ 003383，用来表示上述 VLDLR 的序列，在文献中引用这个序列时，应该以此编号为准。

(4) KEYWORDS（与序列相关的关键词）字段　由该序列的提交者提供，包括该序列的基因产物以及其他相关信息。关键词是一个历史遗留产物，以前在很多情况下被误用。给一个记录加上关键词通常并不十分有效，因为过去有许多序列提交者选用了不在受控词表中的词，并且在整个数据库中用法也不一致。因此，NCBI 不鼓励使用关键词。像本例一样，目前很多序列提交者都不提交任何关键词。

(5) SOURCE（序列来源的物种名）字段　说明该序列是从什么生物体、什么组织得到的，如本例中人源（HOMO SAPIENS）。次关键字 ORGANISM（种属）指出该生物体的分类学地位，如本例人源真核细胞（HOMO SAPIENS EUKARYOTA）等。

(6) REFERENCE（相关文献编号，或递交序列的注册信息）字段　说明该序列中的相关文献，包括 AUTHORS（相关文献作者，或递交序列的作者）、TITLE（相关文献题目）及 JOURNAL（相关文献刊物杂志名，或递交序列的作者单位）等，以次关键词列出。该字段中还列出医学文献摘要数据库 MEDLINE 的代码。该代码实际上是个超文本链接，点击它可以直接调用上述文献摘要。一个序列可以有多篇文献，以不同序号表示，并给出该序列中的哪一部分与文献有关。

(7) FEATURES（序列特性表起始）　具有特定的格式，用来详细描述序列特性。特性表中带有“/db—xref/”标志的字符可以连接到其他数据库，如本例中的分类数据库（taxon：9606）。序列中各部分的位置都在表中标明，蛋白编码区（392～3013）。

接下来是碱基含量字段，给出序列中的碱基组成，如本例中 8465 个 A、5941 个 C、6185 个 G、9134 个 T。ORIGIN 行是序列的引导行，接下来便是碱基序列，以双斜杠行“//”结束。

2. EMBL 数据库

与在 NCBI 的网站上检索 GenBank 数据库一样，EMBL 数据库（http://www.ebi.ac.uk/embl/）可以在 EBI 的网站上检索，如图 14-4。EMBL 数据库的基本单位也是序列条目，

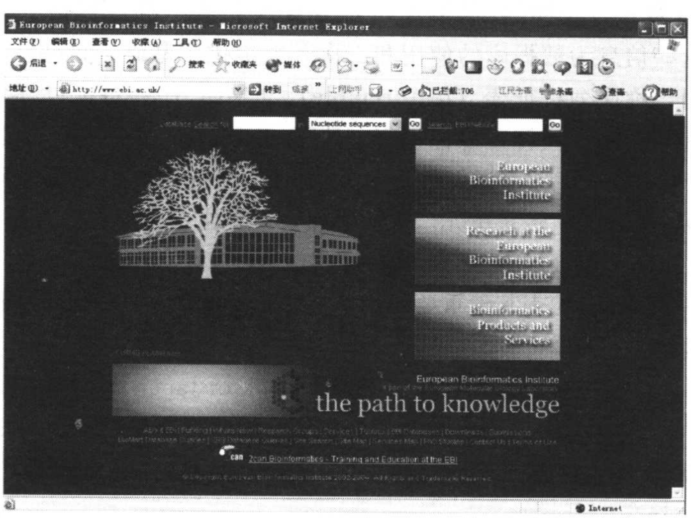


图 14-4　EBI 网站

包括核苷酸碱基排列顺序和注释两部分。序列条目由字段组成，每个字段由标识字起始，后面为该字段的具体说明。有些字段又分若干次子字段，以次标识字或特性表说明符开始，最后以双斜杠“//”作本序列条目结束标记。EMBL 检索界面如图 14-5 所示。

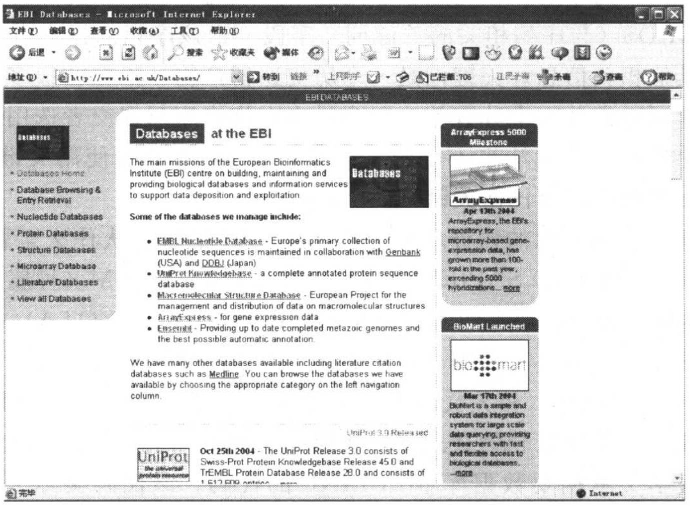


图 14-5　EMBL 检索界面

EMBL 和 GenBank 无论是注释信息，还是序列本身，两者的内容完全相同，只是格式有所区别而已。实际进行数据库检索时，只需检索其中的一个即可（表）。EMBL 条目的关键字包括 ID（序列名称）、DE（序列简单说明）、AC（序列编号）、SV（序列版本号）、KW（与序列相关的关键词）、OS（序列来源的物种名）、OC（序列来源的物种学名和分类学位置）、RN（相关文献编号或递交序列的注册信息）、RA（相关文献作者或递交序列的作者）、RT（相关文献题目）、RL（相关文献杂志名或递交序列的作者单位）、RX（相关文献 Medline 引文代码）、RC（相关文献注释）、RP（相关文献其他注释）、CC（关于序列的注释信息）。DR（相关数据交叉引用号）、FH（序列特征表起始）、FT（序列特征表子项）、SQ（碱基种类统计数）。EMBL 数据格式和 GenBank 格式一样，包括了大量与序列直接相关的注释信息，这些信息为数据库的使用和二次开发提供了基础。这些注释信息位于其他注释信息和序列之间，称序列特征表（featuretable）。EMBL 序列特征以标识字“FH”引导，不同的特征表具有不同的说明符，以标识字“FT”开始。而 GenBank 的特征表则以标识字“FEATURE”引导。序列特征表详细描述该序列的各种特性，包括蛋白质编码区以及翻译所得的氨基酸序列、编码区、转录单元、重复区域、突变位点或修饰位点等信息以及与蛋白质数据库 Swiss-Prot 和分类学数据库 Taxonomy 等其他数据库的交叉索引编号。如 RBS 代表核糖体结合位点，Misc_feature 代表其他重要生物功能区，Allele 代表等位基因的不同形式。

3. DDBJ 数据库

DDBJ 数据库是日本核酸数据库（DNA Data Bank of Japan），由日本国立遗传学研究所遗传信息中心维护。首先反映日本所产生的 DNA 数据，同时与 GenBank、EMBL 合作，互通有无，同步更新，每年四版。该数据库采用与 GenBank 一致的格式，网址为：http://www.ddbj.nig.ac.jp/。

4. BioSino 数据库

BioSino 数据库是中国自主开发的核酸序列公共数据库，发表我国各基因研究中心提供的核酸序列，并接受我国核酸序列的注册登记，由中国科学院上海生命科学研究院生物信息中心维护，其网址为：http://www.biosino.org/。

5. 基因组数据库

一级数据库数据的另一大类来源是众多的基因组测序计划所获得的大量基因组序列数据。目前已经完成了以下基因组测序：代表性的真核生物酿酒酵母、线虫、果蝇。继 2000 年 6 月 26 日科学家公布人类基因组“工作框架图”之后，美、英、日、德、法、中 6 国科学家和美国塞莱拉公司 2001 年 2 月 12 日联合公布了人类基因组图谱及初步分析结果。这次公布的人类基因组图谱是在原“工作框架图”的基础上，经过整理、分类和排列后得到的，它更加准确、清晰、完整。对它的初步分析表明，人类基因组由 31.647 亿个碱基对组成，共有 3 万～3.5 万个基因。

6. GDB 数据库

GDB 数据库是人类基因图谱数据库，目的在于支持构建人类基因图谱和测序。现其主节点移至加拿大多伦多儿童医院生物信息超级计算中心。目前，GDB 含有 150 多万个各种各样的克隆，包括以下几方面的内容：①人类基因组区域，包括基因、克隆、断裂位点、细胞遗传学标记、脆性位点、重复序列等；②人类基因组作图，包括细胞遗传学作图、连锁作图、RH 图、完整作图等；③人类基因组变异，包括突变和多态性以及等位频率数据等。随着人类基因组计划向后基因组计划的转移，GDB 以后还将进一步包括有关功能和表型方面的数据。以上作图数据可调出来逐个图解显示或与其他同源区一起并列图解显示。通过该数据库可直观地了解所需基因的位置信息。可通过基因符号、GenBank 注册号或关键词等进行搜索。其网址为 http：//www.gdb.org/。

7. GeneCards 数据库

GeneCards 数据库是由以色列魏茨曼科学研究所维护的关于基因及其产物以及生物医学应用的文献库。GeneCards 以卡片的形式给出结果，其中列出所查询基因的以下内容：①正式名称、GDB 同义列表、小鼠中的同源物、细胞遗传学定位、基因产物名称、产物功能，如在细胞中的作用、表达方式、定位、与其他蛋白质的同源性及其在疾病中的作用等；②相关基因家族；③相关疾病列表；④有关的研究论文；⑤医学应用，如根据该基因的有关知识而建立的新的治疗与诊断方法等。与 GeneCards 链接的网站有 HUGO（Human Gene Nomenclature Committee）、GDB（Genome Database）、MGD（Mouse Genome Database）等。可通过关键词、基因符号（gene symbol）、基因接受号等进行查询。其网址为 http://nciarray.nci.nih.gov/cards/index.html。

8. UniGene

人类基因组计划的首要任务是对人类基因组进行全序列测定，整个基因组估计有 30 亿个碱基对，其中大约 3%可以编码蛋白质，其余部分的生物学功能还不清楚。转录图谱可以把基因组中能够编码蛋白质的部分集中起来，因此是一种重要的数据资源。UniGene（http://www.ncbi.nlm.nih.gov/UniGene/）试图通过计算机程序对 GenBank 中的序列数据进行适当处理，剔除冗余部分，将同一基因的序列，包括 EST 序列片段搜集到一起，以便研究基因的转录图谱。UniGene 除了包括人的基因外，也包括小鼠、大鼠等其他模式生物的基因。该数据库的标题行（TITLE）给出基因的名称和简要说明，表达部位行（EXPRESS）指出该基因在什么组织中表达以及在基因图谱中的位置等。此外，列出该基因在核酸序列数

据库 GenBank 或 EMBL 和蛋白质序列数据库 SWISS-PROT 编号的超文本链接。

UniGene 中部分条目包括已知基因序列，而有些条目则仅有新测得的 EST 序列片段。这就意味着，这些 EST 序列所对应的基因尚未搞清，可以用来发现新基因。在描绘基因图谱及大规模基因表达分析等研究中，UniGene 也可以帮助实验设计者选择试剂。UniGene 可以通过 NCBI 或 SRS 系统访问。

9. SGD

酵母基因组数据库 SGD（http：//genome-www. stanford. edu/Saccharomyces/）是已经完成基因组全序列测定的啤酒酵母基因组数据库资源，包括啤酒酵母的分子生物学及遗传学等大量信息。通过互联网可以访问该数据库的全基因组信息资源，包括基因及其产物、一些突变体的表型，以及各种有关的注释信息。酵母基因组是于 1998 年完成基因组全序列测定的第一个真核生物基因组。SGD 将各种功能集成在一起，生物学家可通过该数据库进行序列的同源性搜索、对基因序列进行分析、注册酵母基因名称、查看基因组的各类图谱、显示蛋白质分子的三维结构、设计能够有效克隆酵母基因的引物序列等。该数据库通过方便实用、形象生动的图形界面为用户提供酵母基因组的物理图谱、遗传图谱和序列特性图谱等信息。

第三节　常用蛋白质序列数据库

目前常用的蛋白质序列数据库主要有 SWISS-PROT 数据库和 PIR 蛋白质序列数据库。该两大蛋白质序列数据库随同 EMBL 和 GenBank 的增长一起发展。

由于蛋白质序列测定技术先于 DNA 序列测定技术问世。蛋白质序列的搜集也早于 DNA 序列。蛋白质序列数据库的雏形可以追溯到 20 世纪 60 年代。20 世纪 60 年代中期到 80 年代初，美国国家生物医学研究基金会（National Biomedical Research Foundation，简称 NBRF）Dayhoff 领导的研究组将搜集到的蛋白质序列和结构信息以“蛋白质序列和结构图表集”（atlas of protein sequence and structure）的形式发表，主要用来研究蛋白质的进化关系。1984 年，“蛋白质信息资源”（protein information resource，简称 PIR）计划正式启动，蛋白质序列数据库 PIR 也因此而诞生。与核酸序列数据库的国际合作相呼应，1988 年美国的 NBRF、日本的国际蛋白质信息数据库（Japanese international protein information data base，简称 JIPID）和德国的慕尼黑蛋白质序列信息中心（munich information center for protein sequences，简称 MIPS）合作成立了国际蛋白质信息中心（PIR-international），共同收集和维护蛋白质序列数据库 PIR。

除了 PIR 外，另一个重要的蛋白质序列数据库则是 SWISS-PROT。该数据库由瑞士日内瓦大学于 1986 年创建，目前由瑞士生物信息学研究所（Swiss Institute of Bioinformatics，简称 SIB）和欧洲生物信息学研究所 EBI 共同维护和管理。瑞士生物信息研究所下属的蛋白质分析专家系统（expert protein analysis system，简称 ExPASy）的 Web 服务器除了开发和维护 SWISS-PROT 数据库外，也是国际上蛋白质组和蛋白质分子模型研究的中心，为用户提供大量蛋白质信息资源。

由于 DNA 克隆和测序比蛋白质纯化和序列测定容易，大多数蛋白质序列是从编码蛋白质的基因序列翻译而来，因此核酸序列数据库和蛋白质序列数据库有着密切的合作关系。例如，EMBL 和 GenBank 不断将新的编码蛋白质的基因序列传送到 PIR，而 SWISS-PROT 数据库的记录每天由新的 EMBL 记录翻译过来。

1. SWISS-PROT 蛋白质序列数据库

SWISS-PROT（http：//www.ebi.ac.uk/swissprot/）是蛋白质序列数据库，该数据库力图提供高水平的数据注释信息，包括对蛋白质功能、结构域的结构、蛋白质翻译后的修饰、突变体等的描述。SWISS-PROT 数据库的目标是提供尽可能详尽的、很少冗余的数据，数据库中的链接可以直接链接到其他资源上。

SWISS-PROT 蛋白质序列数据库由大量序列条目组成，每一个序列条目有其自己的格式。为了标准化的目的，SWISS-PROT 的格式与 EMBL 核酸序列数据库的格式尽可能类似。SWISS-PROT 涉及已知蛋白质的功能、序列（包括一些蛋白质片段的序列）、结构域（如跨膜区等）的结构、翻译后修饰（如磷酸化与去磷酸化等）及其位点、突变体等，目前已经收录了近 8 万个序列。通过它可以查询所需蛋白质的上述所有内容，从而指导进一步的实验。SWISS-PROT 还与其他一些数据库相链接，包括 PROSITE、SWISS-2DPAGE、SWISS-3D IMAGE、ENZYME、SWISS-MODEL、NCBI 的数据库等，因而可进行更为广泛的查询工作。SWISS-PROT 可通过关键词、SWISS-PROT 接受号、作者等进行查询。记录的题头缩写见表 14-1。

表 14-1　SWISS-PROT 数据库标记的题头缩写

缩写	全　称	含　义	缩写	全　称	含　义
ID	Identification	鉴别号	RC	Reference comments	参考文献评注
AC	Accession number	存取号	RX	cross-references	参考文献的 MEDLINE 号
DT	Date	注录日期	RA	Reference authors	参考文献作者
DE	Description	名称	RL	Reference location	参考文献出处
GN	Gene name(s)	基因名称	CC	Comments or notes	评注
OS	Organism species	来源物种	DR	Database cross-references	其他数据库指针
OG	Organelle	来源细胞器	KW	Keywords	关键词
OC	Organnism classification	物种分类	FT	Feature table data	性质列表
RN	Reference number	参考文献号	SQ	Sequence header	序列题头
RP	Reference position	参考文献涉及内容	//	Termination line	结束符

2. PIR 蛋白质序列数据库

PIR 蛋白质序列数据库（http：//pir.georgetown.edu/）是由美国国家生物医学研究基金会 NBRF 的 Margaret Dayhoff 在 20 世纪 60 年代早期发展建立的，Margaret Dayhoff 收集蛋白质序列来研究蛋白质的进化关系。自从 1988 年，蛋白质序列数据库就由国际蛋白质信息中心 PIR-International 的成员单位来共同维护，该数据库所收集的内容包括了 NBRF 蛋白质信息资源（PIR）、日本的国际蛋白质信息数据库（JIPID）和 Martinsried 研究所的蛋白质序列（MIPS）。PIR 是国际上最大的公共蛋白质序列数据库，它是一个全面的、经过注释的、非冗余的蛋白质序列数据库，包含超过 283 366 条蛋白质序列（至 2003 年 11 月），其中包括来自几十个完整基因组的蛋白质序列。所有序列数据都经过整理，超过 99%的序列已按蛋白质家族分类，一半以上还按蛋白质超家族进行了分类。PIR 数据库按照数据的性质和注释的层次分为 4 个不同的部分，即 PIR1～PIR4，见表 14-2。

表 14-2 PIR 数据库的组成

名 称	描 述	说 明
PIR1	分类和注释的序列	完全分类清楚
PIR2	注释的序列	已检查和分类
PIR3	未确证的序列	未检查
PIR4	非自然存在的序列	未解码翻译

PIR 数据库实际上是蛋白质序列数据库 PSD（protein sequence database）、PATCHX（收入尚未纳入 PIR 库的蛋白质序列）、ARCHIVE（保存 PIR 库中条目的原始文献或最初提交的序列）、NRL-3D、FAMBASE、PIR-ALN、RESID、ProCLASS（根据 PROSITE 库和 PIR 库中超家族的关系组织起来的非冗余蛋白质库）、ProtFam 和 PIR-ASDB（PIR 的注释和相似性数据库）等多个数据库的集成和链接，通过提供蛋白质序列数据库、衍生的相关数据库以及相应的软件而支持有关分子进化、功能基因组学和计算生物学方面的研究，包括：①能够快速查询、比较蛋白质序列并对其进行特征序列的模式匹配；②可给出蛋白质的功能位点，如磷酸化位点、糖基化位点、细胞黏附位点、与其他蛋白质的共有序列等；③可进行多种方式的序列比对，如数据库相似性搜索、双序列比对和多序列比对等。可通过关键词、特征序列或序列序号等进行查询。图 14-6 为 PIR 数据库的检索界面。

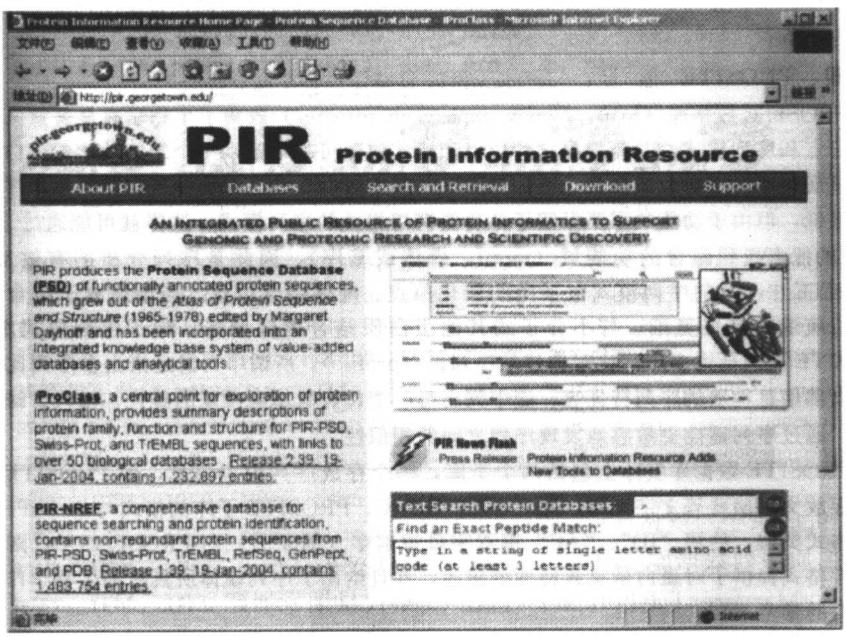


图 14-6 PIR 数据库检索界面

3. 其他蛋白质序列数据库

PROSITE 数据库（http://www.expasy.ch/prosite/）收集了生物学有显著意义的蛋白质位点和序列模式，并能根据这些位点和模式快速可靠地鉴别一个未知功能的蛋白质序列应该属于哪一个蛋白质家族。有的情况下，某个蛋白质与已知功能蛋白质的整体序列相似性很低，但由于功能的需要保留了与功能密切相关的序列模式，这样就可能通过 PROSITE 的搜索找到隐含的功能域（motif，模体或基序），因此是序列分析的有效工具。PROSITE 中

涉及的序列模式包括酶的催化位点、配体结合位点、与金属离子结合的残基、二硫键的半胱氨酸、与小分子或其他蛋白质结合的区域等。除了序列模式之外，PROSITE 还包括由多序列比对构建的序列谱（profile），所谓序列谱是指由多序列比对结果的全部信息构造的序列特征表，表中列出每一个位点氨基酸残基的替代、插入、缺失等情况，通过序列谱能更敏感地发现序列之间的相似性。

TrEMBL（翻译的 EMBL）数据库（http：//www.ebi.ac.uk/trembl/index.html）是1996 年创建并增补到 SwissProt 数据库中。该数据库采用 SwissProt 数据库的格式，包含有对 EMBL 数据库中所有编码序列的翻译。TrEMBL 数据库有两个主要部分，SP-TrEMBL 的条目已由专家人工分类并且赋予了 SWISS-PROT 库的索取号，但是还没有通过人工审读并最终收入 SWISS-PROT；REM-TrEMBL（REMaining TrEMBL）包含由于某种原因而还没有被收入 SWISS-PROT 的条目，主要是一些免疫球蛋白和 T 细胞受体、少于 8 个氨基酸的碎片、合成的序列、具有专利的序列和不能编码真正蛋白质的密码子。TrEMBL 数据库的结构与 SwissProt 数据库一样，具有很好的可查询功能，而且不会由于组合到 SwissProt 数据库中的序列没有被充分分析和注释而影响到 SWISS-PROT 数据库。

此外还有基于蛋白质指纹技术的数据库 PRINTS 数据库（http：//www.bioinf.man.ac.uk/dbbrowser/prints/），及从蛋白质结构数据库 PDB 中提取出来的序列构成的数据库 NRL-3D（http：//nbrfa.georgetown.edu/pir/nrl3d.html），在此不一一详述。

4. 蛋白质结构数据库

除了生物序列数据库外，生物大分子三维空间结构数据库则是另一类重要的分子生物信息数据库。根据分子生物学中心法则，DNA 序列是遗传信息的携带者，而蛋白质分子则是主要的生物大分子功能单元。蛋白质分子的各种功能，是通过不同的三维空间结构实现的。因此，蛋白质空间结构数据库是生物大分子结构数据库的主要组成部分。蛋白质结构数据库是基于 X 射线晶体衍射和核磁共振分析所获得的晶体结构数据，其基本内容为实验测定的蛋白质分子空间结构原子坐标。20 世纪 90 年代以来，越来越多的蛋白质分子结构被测定，蛋白质结构分类的研究不断深入，出现了蛋白质家族、折叠模式、结构域、回环等结构分类数据库。虽然目前其中的序列数量远比不上蛋白质序列数据库，但其数据量以指数形式迅速增加。

蛋白质结构数据库（protein data bank，PDB）是在 20 世纪 70 年代由美国自然科学基金会、能源部和国立卫生研究院共同投资建立，主要由 X 射线晶体衍射和核磁共振（NMR）测得的生物大分子三维结构所组成，用户可直接查询、调用和观察库中所收录的任何大分子三维结构。该数据库同时提供蛋白质序列及其三维空间晶体学原子坐标，其中受体-配体、抗原-抗体、底物-酶复合物等相互作用分子的共结晶图谱是基于同源比较的分子设计所需的最佳模型，因此 PDB 数据库为初步的蛋白质合理设计提供了重要的知识来源。由于 PDB 主要由生物大分子三维结构所组成，它具有以下几种功能：①能够查找目的蛋白质的结构；②可进行一级或高级结构的简单分析；③与互联网上的其他一些数据库，如 GDB、GenBank、SWISS-PROT、PIR 等链接，从而可查询蛋白质的其他信息；④可下载有关的结构信息以供进一步使用。可通过关键词或 PDB 标识符等进行查询。在序列分析中，PDB 主要可应用于蛋白质结构预测和结构同源性比较。其中前面提到的 NRL-3D 数据库则是 PDB 数据库中所有蛋白质序列的信息。该数据库允许进行基于结构的序列比较。PDB 数据库的检索界面见图 14-7。

在过去，生物大分子的成像是不可能的。1972 年 Levinthal 和 Katz 开创性地建立了计



图 14-7 PDB 数据库的检索界面

算机线框（wireframe）模型。几年以后，人们利用计算机完成立体分子及其旋转模型的愿望实现了。近年来，已有一些方便小巧的如 Kinemage 和 Rasmol 的三维结构浏览软件可使我们直观地了解蛋白质结构。利用 PDB 数据库中蛋白质的结构信息和相关的软件，我们可以方便地从各种角度观察蛋白质的空间结构（图 14-8）。

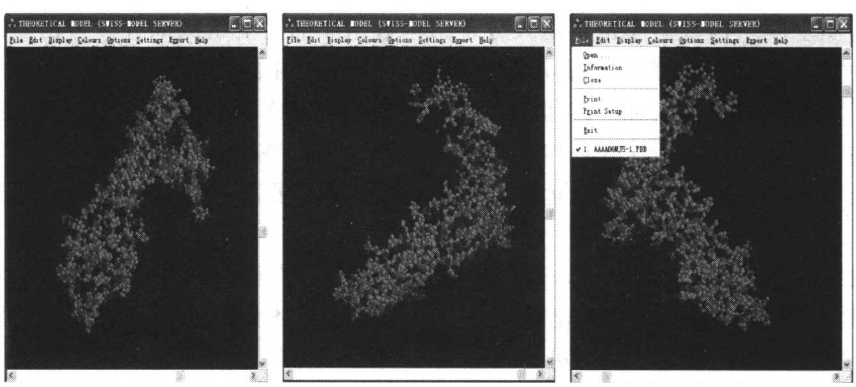


图 14-8 用 Rasmol 显示蛋白质的结构（不同视角）

第四节 序 列 分 析

在生物学的研究中，一个最常用和最经典的研究手段，就是通过比较分析获取有用的信息和知识。通过将研究对象相互比较来寻找对象可能具备的特性。我们从核酸以及氨基酸的

一级结构分析序列的相同点和不同点，以期能够推测它们的结构、功能以及进化上的联系。最常用的比较方法是序列对比，它为两个或更多个序列的残基之间的相互关系提供了一个非常明确的图谱。通过比较两个序列之间的相似区域和保守性位点，寻找二者可能的分子进化关系。进一步的对比是将多个蛋白质或核酸同时进行比较，寻找这些有进化关系的序列之间共同的保守区域、位点和图谱，分析产生共同功能的序列模式。此外，还可以把蛋白质序列与核酸序列相比来探索核酸序列可能的表达框架；把蛋白质序列与具有三维结构信息的蛋白质相比，从而获得蛋白质空间结构的信息。

20 世纪 70 年代以来，DNA 测序方法的飞速发展，极大地引发了序列信息量的扩增，从而使可供比较的序列数量呈现爆炸式增长。分子生物学家意识到，将未知序列同整个数据库中的已知序列进行比较分析已经成为一个强有力的研究手段。在过去的 30 年里，序列比较的各种算法发展得越来越迅速，也越来越成熟，已经能够跟上序列数据库增长的步伐。通过序列比较能快速地获得有关查询序列的大量有价值的参考信息，对于进一步分析其结构和功能都会有很大的帮助。近年来随着生物信息学数据大量积累和生物学知识的整理，通过对比方法可以有效地分析和预测一些新发现基因的功能。

一、序列比对方法和相似性搜索

生物物种之间存在进化关系，对基因和蛋白质序列进行比较，从本质上来讲，是进行同达尔文进化论一样的比较分析，只不过更加精细、更加详尽；如果两个序列之间具有足够的相似性，就推测二者可能有共同的进化祖先，经过序列内残基的替换、残基或序列片段的缺失以及序列重组等遗传变异过程分别演化而来。值得注意的是序列相似和序列同源是不同的概念，序列之间的相似程度是可以量化的参数，而序列是否同源需要有进化事实的验证。

大多数对比方法很自然地都希望能够在某种程度上建立起分子进化的模型。我们通常都假定同源序列是从某一共同祖先不断变化而来，祖先序列在进化过程中发生取代、插入以及缺失的变化。

在序列比较过程中，对同源基因或蛋白质序列相互比较时，残基之间相互对应，使取代的情况很明显地表现出来。

在残基-残基对比中，很明显，某些位置的氨基酸残基相对于其他位置的残基具有较高的保守性，这个信息揭示了某些残基对于一个蛋白质的结构和功能是极为重要的。一般说来，处于活性位点的残基都是极为保守的，比如形成二硫键的半胱氨酸、参与电子传递的氨基酸残基以及决定底物特异性的氨基酸残基。这些保守的残基对于保持蛋白的结构与功能非常重要，另一方面，由于历史原因，某些保守位置对蛋白功能并无太大的重要性。当我们处理非常相近的物种时必须十分小心，因为相似性在某些情况下更多的是历史的反映，而不是功能的反映，比如，小鼠和大鼠的某些序列具有高度的相似性，可能仅仅是因为没有足够的时间进行分化而已。尽管如此，序列对比仍然是从已知获得未知的一个十分有用的方法，比如通过比较一个新的蛋白同其他已经深入研究的蛋白，可以推断这个未知蛋白结构与功能的某些性质，例如它们可能是酶的活性位点残基、形成二硫键的半胱氨酸残基、与配体结合部位的残基及与金属离子结合的残基、形成特定结构基序（motif）的残基等。但并不是所有保守的残基都一定是结构功能重要的残基，可能它们只是由于历史的原因被保留下来，而不是由于进化压力而保留下来。因此，如果两个序列有显著的保守性，要确定二者具有共同的进化历史，进而认为二者有近似的结构和功能，还需要更多实验和信息的支持。通过大量实

验和序列对比的分析，一般认为蛋白质的结构和功能比序列具有更大的保守性，因此粗略地说，如果序列之间的相似性超过 30%，它们就很可能是同源的。必须指出的是，理论分析只提供了序列进化的可能性，不能够仅仅是通过比较分析这一论据来断定结论是否正确，结论还必须经过实验验证。

早期的序列对比是全局的序列比较，但由于蛋白质具有的模块性质，可能产生由于外显子交换而产生新蛋白质，因此局部对比会更加合理。通常用打分矩阵描述序列两两对比，两条序列分别作为矩阵的两维，矩阵点是两维上对应两个残基的相似性分数，分数越高则说明两个残基越相似。因此，序列对比问题变成在矩阵里寻找最佳对比路径，目前最有效的方法是 Needleman-Wunsch 动态规划算法，在此基础上又改良产生了 Smith-Waterman 算法和 SIM 算法。在 FASTA 程序包中可以找到用动态规划算法进行序列对比的工具 LALIGN，它能给出多个不相互交叉的最佳对比结果。

在进行序列两两对比时，有两方面问题直接影响相似性分值：取代矩阵和空位罚分。粗糙的对比方法仅仅用相同和（或）不同来描述两个残基的关系，显然这种方法无法描述残基取代对结构和功能的不同影响效果，缬氨酸对异亮氨酸的取代与谷氨酸对异亮氨酸的取代应该给予不同的打分。因此，如果用一个取代矩阵来描述氨基酸残基两两取代的分值，会大大提高对比的敏感性和生物学意义。虽然针对不同的研究目标和对象应该构建适宜的取代矩阵，但国际上常用的取代矩阵有 PAM 和 BLOSUM 等，它们来源于不同的构建方法和不同的参数选择，包括 PAM250、BLOSUM62、BLOSUM90、BLOSUM30 等。对于不同的对象可以采用不同的取代矩阵以获得更多信息，例如对同源性较高的序列可以采用 BLOSUM90 矩阵，而对同源性较低的序列可采用 BLOSUM30 矩阵。

空位罚分是为了补偿插入和缺失对序列相似性的影响，由于没有什么合适的理论模型能很好地描述空位问题，因此空位罚分缺乏理论依据而更多地带有主观特色。一般的处理方法是用两个罚分值，一个对插入的第一个空位罚分，如 10～15；另一个对空位的延伸罚分，如 1～2。对于具体的对比问题，采用不同的罚分方法会取得不同的效果。

对于对比计算产生的分值，到底多大才能说明两个序列是同源的，对此有统计学方法加以说明，主要的思想是把具有相同长度的随机序列进行对比，把分值与最初的对比分值相比，看看对比结果是否具有显著性。相关的参数 E 代表随机对比分值不低于实际对比分值的概率。对于严格的对比，E 值必须低于一定阈值才能说明对比的结果具有足够的统计学显著性，这样就排除了由于偶然的因素产生高对比得分的可能。

GenBank、SWISS-PROT 等序列数据库提供的序列搜索服务都是以序列两两对比为基础的，不同之处在于为了提高搜索的速度和效率，通常的序列搜索算法都进行了一定程度的优化，如最常见的 FASTA 工具和 BLAST 工具。

二、FASTA

FASTA 是第一个被广泛应用的序列对比和搜索工具包，包含若干个独立的程序。FASTA 为了提高序列搜索的速度，会先建立序列片段的“字典”，查询序列先会在字典里搜索可能的匹配序列，字典中的序列长度由 ktup 参数控制，缺省的 ktup＝2。FASTA 的结果报告中会给出每个搜索到的序列与查询序列的最佳对比结果，以及这个对比的统计学显著性评估 E 值。FASTA 工具包可以在大多提供下载服务的生物信息学站点上找到。也可在 http：//www. ebi. ac. uk/fasta/进行在线检索和比对，图 14-9 为其搜索界面。

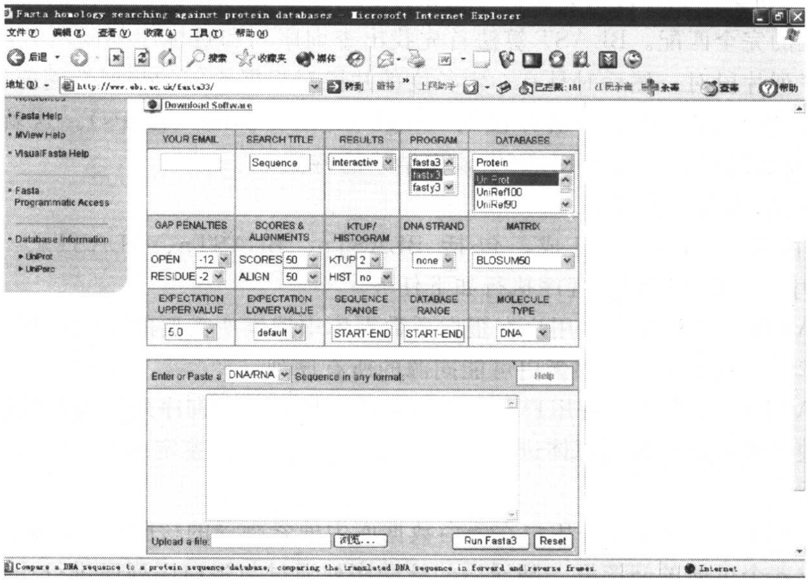


图 14-9　FASTA 搜索界面

将待检索的序列粘贴到指定部位或以文件上传后，在界面中选定相关参数，如比对的类型（核酸或蛋白质）、所用数据库、所用程序类别、比对方法、ktup 值（1 或 2）、数据返回模式（立即以网页形式返回或以 E-mail 返回），点击 RUN FASTA，即可得到比对结果。由 FASTA 搜索产生的典型输出结果包括一般信息和数据库搜索结果。一般信息包括所用搜索程序名称、版本号、该程序的文献出处、所提交的待检序列以及数据库名称和版本号，程序运行所用参数和运行时间等。紧跟这些一般信息的是数据库搜索结果：首先列出搜索得到的目标序列简单说明，其数目可由用户定义。所列出目标序列的信息包括序列所在数据库名称缩写（如 SWALL 表示 Swiss Prot All 库）、目标序列的识别符 ID、序列编号 AC 和序列名等部分信息。括号中标明匹配部分的残基数；紧接着是由程序计算得到的初始化和优化后的分数值；最后一列是期望值 E，用来判断比对结果的置信度。接近于 0 的 E 值表明两序列的匹配不大可能是由随机因素造成的；在其前一列的为 bits 值，其值越大，E 值越小，反映序列的相似性越高。

在对搜索结果给出概要信息后，输出结果中详细列出待检序列与目标序列比对的详细结果，其数目也可以由用户选择。为减少输出篇幅，有时在第一次运行时只输出概要信息和部分详细比对结果。比对结果给出待检序列和目标序列间匹配残基的排列情况，完全匹配用冒号“:”表示，相似性匹配用点“.”表示。残基相似性定义取决于所用相似性分数矩阵，也可以由用户选择。输出结果的概要信息部分列出所用相似性矩阵的名称。比对结果上方数字表示待检序列残基编号，下方数字表示目标序列残基编号，此处均不计插入的空位。

三、BLAST

BLAST（basic local alignment search tool）（http://www.ncbi.nlm.nih.gov/BLAST/）是现在应用最广泛的序列相似性搜索工具，相比 FASTA 有更多改进，速度更快，并建立在严格的统计学基础之上。BLAST 算法的基本要点是序列片段对（segment

pair）的概念。所谓序列片段对是指两个给定序列中的一对子序列，它们的长度相等，且可以形成无空位的完全匹配。BLAST 算法首先找出查询序列和目标序列间所有匹配程度超过一定阈值的序列片段对，然后对具有一定长度的片段对根据给定的相似性阈值延伸，得到一定长度的相似性片段，称高分值片段对（high scoring pairs，HSPs）。这就是无空位的 BLAST 比对算法的基础，也是 BLAST 输出结果的特征。

1. BLAST 的功能及组成

BLAST 由多个子程序组成，包括 BLASTp、BLASTn、BLASTx、tBLASTn 和 tBLASTx 程序，五种 BLAST 程序执行如下任务。

（1）BLASTp　该程序允许用户在蛋白质数据库中搜索所需要的蛋白序列。可用于在查询序列数据库中寻找与某一已知蛋白可能同源的所有序列。

（2）BLASTn　该程序允许用户在核酸数据库中搜索与查询序列同源的核酸序列。一个新测序的核酸序列可以与其同源体进行比较，以对该序列进行鉴定或发现是否有可能有其他序列片段杂质。

（3）BLASTx　该程序允许用户在蛋白数据库中搜索翻译的核酸序列。被查询的核酸序列先被翻译成 6 个可能的阅读框。翻译后的核酸序列通过与蛋白序列库进行比较，发现其中可能的同源性蛋白。也可用于寻找核酸测序中的错误。在这种情况下，该程序尤其有用。BLASTx 输出文件的信息也可以帮助鉴定特定核酸序列中尚不清楚的基因。

（4）tBLASTn　允许用户以蛋白查询序列在特定的核酸数据库中搜索翻译的核酸序列。在特定的核酸数据库中核酸序列一开始被翻译成 6 个可能的阅读框，然后与蛋白查询序列进行比较。在通过将蛋白查询序列与特定的核酸数据库中翻译的核酸同源体进行比较来发现蛋白测序错误时，该程序特别有用。tBLASTn 输出文件中的信息还有助于澄清特定查询序列中不明的氨基酸残基。就 6 个阅读框翻译比较方法而言，tBLASTn 与 BLASTx 相似，但它不是使用核酸查询序列去查询，而用蛋白查询序列。由于该程序需要将特定核酸序列数据库中所有序列都翻译成 6 个可能的读框，所以运行时间较长。

（5）tBLASTx　该程序首先将查询的核酸序列翻译成 6 个可能的读框，然后将核酸序列库中所有序列也翻译为 6 个可能的读框，最后将查询序列的翻译结果与核酸序列数据库的翻译结果进行相似性比较，以发现相似序列。TBLASTx 与 BLASTx 和 tBLASTn 的程序相似，是 BLASTx 搜索的补充。

除了上述常用的 BLAST 系列程序以外，NCBI 还开发了以下多种扩展软件。

Gapped BLAST　BLAST2.0 以前的版本只能匹配连续的序列，如果比对中出现空位，结果将被分段显示。而 Gapped BLAST 允许在序列比对中引入空位，这将更好地复原实际情况。

MEGBLAST　提供大量长序列的比较。

PSI-BLAST　搜索证实远源进化关系的存在与否和进一步获取这个蛋白家族中的功能信息。

PHI-BLAST　搜索蛋白基序。

上述 BLAST 程序在 BLAST 搜索主界面中都可找到（见图 14-10）。

2. BLAST 的检索方法

下面以人 VLDLR 的核酸序列为例来说明 BLAST 的检索方法。因为是在核酸序列数据库中搜索，所以选择 BLASTn，进入如图 14-11 所示界面。

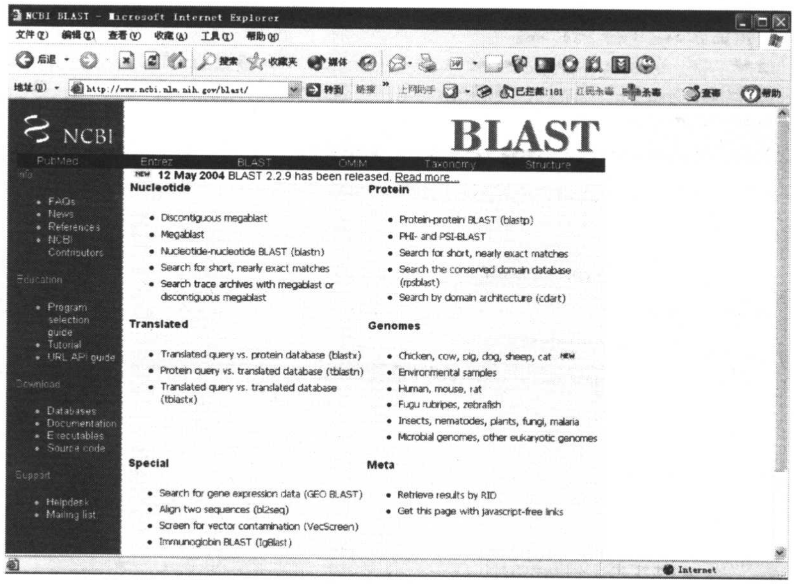


图 14-10　BLAST 搜索的主界面

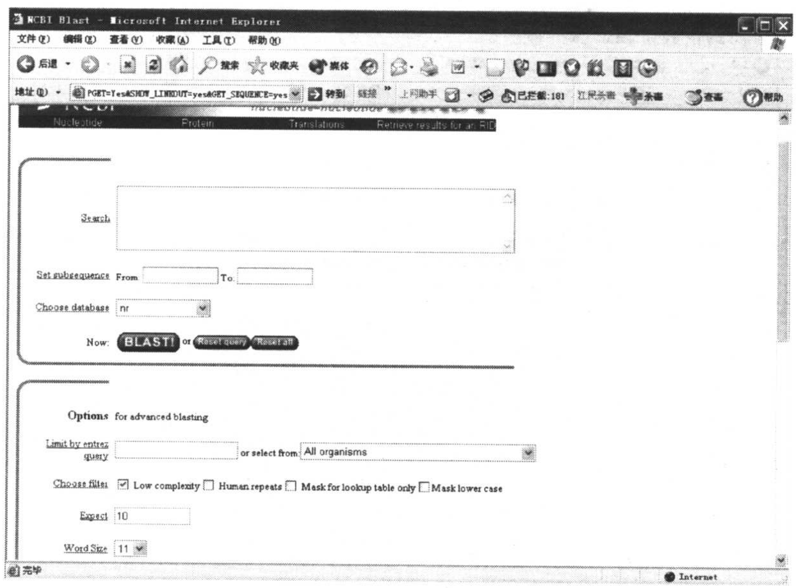


图 14-11　BLASTn 搜索界面

在“Search”文本框中输入 VLDLR 的序列，设定好搜索参数（一般情况下用默认值即可）后即可点击“BLAST!”按钮进行搜索。服务器首先返回一个所提交任务的标志（RID）号（见图 14-12），用户通过点击“Format results”后可等待传输分析结果。

返回的结果由三部分组成。

（1）结果总览图　用红、粉、绿、蓝、黑五种颜色代表同源程度（图 14-13）。红色最高，排在最上面。其他各种颜色同源性逐渐降低，每条图代表搜索蛋白质匹配的序列。如果出现阴影区，其对应的是两个或多个搜索数据库相似序列中的非相似区。鼠标指在哪条线

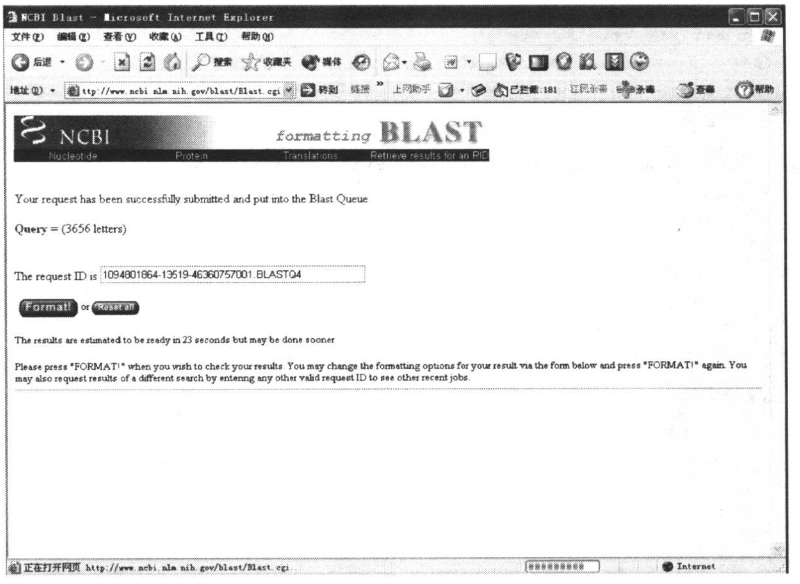


图 14-12　BLASTn 搜索结果确认界面

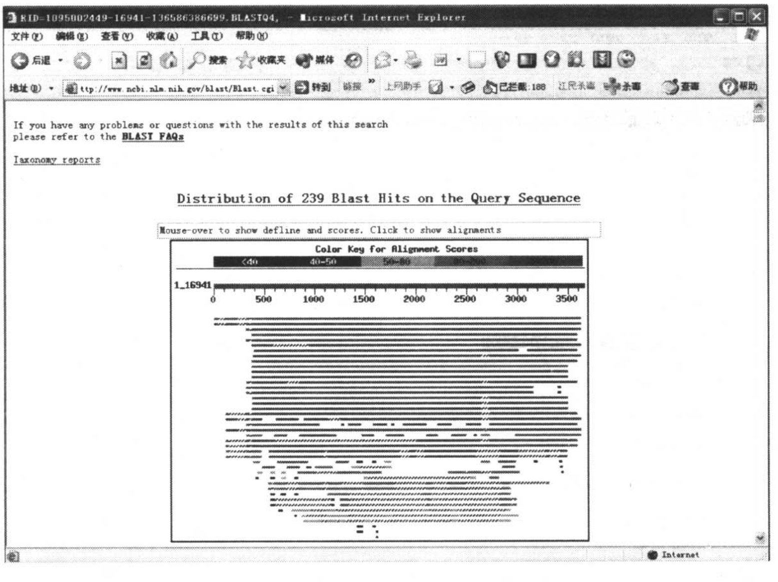


图 14-13　BLASTn 搜索结果总览图

上，最上面的文本框就会显示匹配蛋白质的名称，同一线条内的分离部分对应于无关的采样数。

(2) 显著性序列列表　如图 14-14 所示，具有显著性（同源性最高）的排行在最上面，其期望值 E 最低，排行按照 E 值增加排序。每一行有四部分内容：①数据序列标识符；②对该序列的简单描述；③在每个数据库中搜索得到的分数；④E 值。点击序列标识符可链接到相应序列库中，点击分数可链接到对应的详细行列比对。

(3) 行列比对　图 14-15 所示的为人 VLDLR 与鼠 VLDLR 序列的比对。系统一般默认的是配对行列输出格式，即查询序列与数据库中匹配的序列垂直对应。在 BLAST 搜索中，

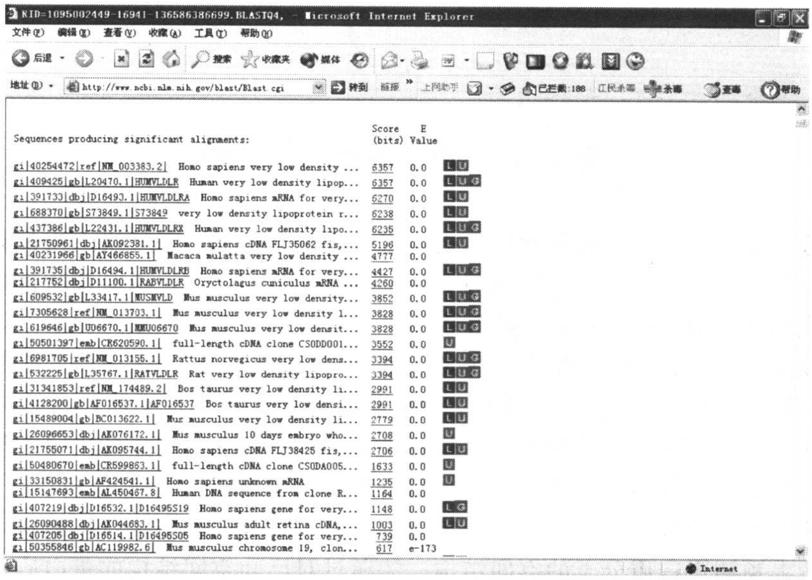


图 14-14　搜索结果的显著性序列列表

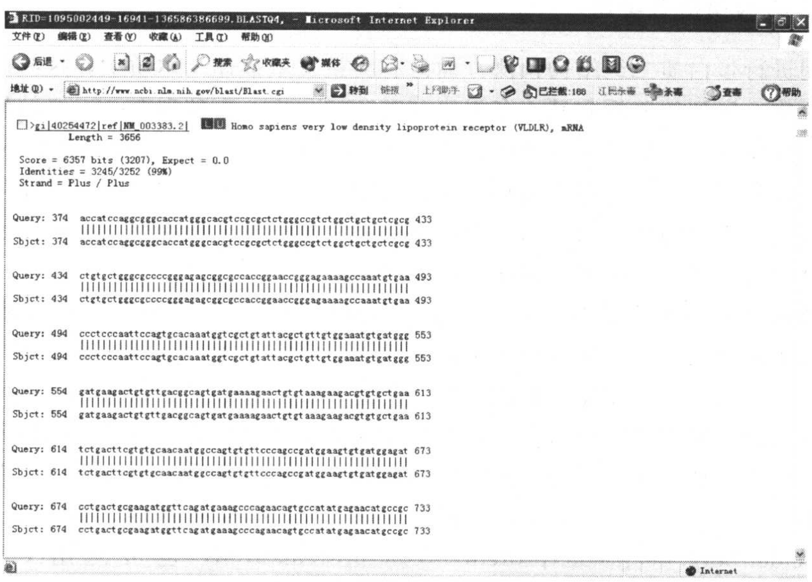


图 14-15　人 VLDLR 与鼠 VLDLR 序列的比对结果

对于蛋白质查询而言，第二行序列显示了查询序列和目标序列间的匹配关系。相同的残基直接列出，保守性残基以“＋”表示，空位部分代表查询序列与目标序列不一致。由于过滤作用，在低复杂区氨基酸查询序列可以包含 Xs（核苷酸包含 Ns）。

相关的参数 E 值代表随机比较分值不低于实际比较分值的概率。对于严格的比较，E 值低于一定阈值才能说明比较的结果具有足够的统计学显著性，这样就排除了由于偶然的因素产生高比较得分的可能。具体界定值在搜索设置中设定，系统默认为 10，搜索的严谨度越高，E 值越小。若要查询短序列，获得更多的信息，则可增加 E 值到 1000 或更高。

第五节　蛋白质结构预测

蛋白质分子是许多氨基酸通过肽键相连形成的生物大分子，每种蛋白质都有其一定的氨基酸组成、氨基酸排列顺序以及肽链空间的特定排布位置。因此，由氨基酸排列顺序及空间排布等所构成的蛋白质分子结构，才能真正体现蛋白质的个性。组成蛋白质的氨基酸序列为蛋白质的一级结构，蛋白质的一级结构决定了蛋白质的性质。组成蛋白质的氨基酸的物理和化学性质早已被人熟知。构成蛋白质的 20 种氨基酸由于化学结构不同，在结构和功能上具有多样性，任一残基对蛋白质的物理和生化性质都会产生影响，即序列决定构象。由于蛋白质空间结构的基础是一级结构，近年来根据蛋白质的氨基酸序列预测其空间结构受到科学家的关注。预测方法主要有两类：①采用分子力学、分子动力学的方法，根据物理化学的基本原理，从理论上预测蛋白质分子的空间结构；②通过对已知空间结构的蛋白质进行分析，找出一级结构与空间结构的关系，总结出规律，用于新的蛋白质空间结构的预测。

一、二级结构和折叠类型

基于序列进行蛋白质二级结构预测方面已有大量文献描述。本质上，这些研究可被分为两大类：基于单一序列的分析和基于多重序列对齐的分析。直到现在，二级结构预测的成功率仍然不高。即使是最好的算法（例如 Chou-Fasman 和 GOR）也只能达到约 60%的预测准确率，尤其是含有较多 β-片层结构时预测往往失败。该问题主要源于单一的序列信息只能提供每个氨基酸残基周围的残基信息而不能提供空间结构信息。局部二级结构往往是蛋白质的自身折叠信息，而这些无法从单一的序列信息中获得。

两方面的研究进展改变了该状态。首先，多重序列对齐可以提高此种预测方法的准确性。一个蛋白质序列的多重序列对齐可被视为天然的突变结果：分析序列中某个位点的变异可以提供蛋白质三级结构的信息。其次，神经网络算法已经被用于根据序列进行结构分析方面的研究。现已公认根据理想的蛋白质序列多重对齐结果进行蛋白质二级结构预测的成功率大为提高——至少比根据单一序列预测的准确性高出 10%。文献报道 PHD 程序是目前此方面的最好程序，提供了从二级结构到折叠方面分析的多种资源，其网址为 http：//www.embl-heidel-berg.de/predictprotein/predictprotein.html，也可使用 e-mail 服务器(predictprotein@embl-heidelberg.de）进行数据分析。此外，ExPASy 上的 Proteomic Tools 提供了许多二级结构预测工具的链接，如 NNPREDICT（http：//www.cmpharm.ucsf.edu/～nomi/nnp-redict.html)、PredictProtein（http：//maple.bioc.columbia.edu/pp/)、SOPMA3（http：//www.ibcp.fr/pridict.html）等。NNPREDICT 是用于预测氨基酸序列中各个氨基酸对应的二级结构。PredictProtein 程序是一个综合性程序，它的功能很全，既包含数据库搜索部分又包含蛋白预测部分。数据库部分搜索包括多序列比较（Maxhom)、功能区检测（PROSITE)、组成偏好-低复杂区域的检测（SEG)、结构域（PRODOM)、折叠区域（TOPITS)；预测部分包括二级结构、跨膜螺旋、球形蛋白、卷曲螺旋、二硫键、结构转换区。对用户输入的查询序列，程序既可以进行同源性比较，又可以进行二级结构预

测。SOPMA 采用了 5 种方法对蛋白质二级结构进行预测，然后将结果优化组合，汇集整理成一个“一致结果”。5 种方法包括：Garnier-Gibrat-Robson（GOR）方法、Levin 同源预测方法、双重预测方法、PredictProtein 中的 PHD 方法和 SOPMA 方法。SOPMA 用这种方法建立已知二级结构蛋白的数据库，库中的每个蛋白都经过二级结构预测，然后用从库中得到的信息对查询序列进行二级结构预测。

二、三级结构预测

三级结构是蛋白质结构预测的最终目的。由于蛋白质结构尤其是高级结构决定蛋白质的功能，因此要充分研究蛋白质的功能，就要把结构研究清楚。生物信息学发展到现在已经有能力对一个未知结构的蛋白质序列作出一系列分析，最终得出一个三级结构模型，这样就可以大大减少研究人员的工作量，而且对于可信度比较高的预测，如同源模建等，对研究蛋白质的生物学活性中心有很好的指导作用。

1. 同源模建

从氨基酸序列预测三级结构的模拟技术很多，其中之一是同源模建，它是通过比较未知结构的蛋白质序列与已知结构的蛋白质序列来预测蛋白质的结构。同源模建也称比较模建（compatative modeling），是目前应用最成功的一种方法，因为同源蛋白质具有相似的结构和功能，所以利用结构已知的同源蛋白质可以建立目标蛋白质的结构模型，再用理论计算方法进行优化，利用同源模建方法建立的蛋白质是以已知的同源蛋白质结构为基础的，所以这是一种基于现代生物学知识的预测方法。同源模建的基本过程包括 6 个部分：目标序列与模板序列的匹配；根据同源蛋白质的多重序列匹配结果，确定同源蛋白质的结构保守区及相应的框架结构；目标蛋白质结构保守区的主链模建；目标蛋白质结构变异区的主链模建；侧链的安装和优化；对模建结构进行优化和评估，力求得到可以解释生物及生化特性的合理模型，因此序列匹配对建立精确结构模型起着关键的作用。一般地，序列同源性越差，匹配的准确程度越低。序列相似性低于 30%的蛋白质难以得到理想的结构模型。

自 20 世纪 50 年代后期至 90 年代初期，Blumdell 等提出同源蛋白质结构预测模型以后，相继出现大量关于同源预测的成果。目前对于序列相似性大于 30%的蛋白质序列已经能够模建相当精确的结构模型。网上服务器有 Swiss-MODEL、CPHmodel。

2. 折叠识别（fold recognition）

某些蛋白质在结构已知的数据库中找不到序列同源性大于 30%的同源蛋白质，但是许多序列同源性很差的蛋白质却存在相同的框架结构——折叠子（fold），此时就可以成功地应用折叠识别方法。1987 年，Finkelstein 和 Ptitsyn 就指出，由于各种立体化学的限制，蛋白质折叠的数目是有限的。后来许多学者对自然界中可能存在的折叠子数目作了估计，1992 年 Chothia 估计自然界中折叠子不会超过 1000，1998 年我国科学家王志新院士作了更精确的估计，认为仅有 654 种折叠子存在。以结构已知的蛋白质的折叠子为模板，寻找给定氨基酸序列可能采取的折叠类型，即为折叠识别。目前利用折叠识别预测蛋白质结构的主要方法大多是从 1991 年 Bowie 等提出的三维剖面（3D-profile）方法和 1992 年 Jones 等提出的线程（threading）方法发展而来的。

3. 从头预测（ab initio prediction）

无论是同源模建还是折叠识别都需要已知的蛋白质结构作为模板，因此这两种方法都不能得到全新的结构，而且在找不到模板结构的情况下难以得到成功应用。与之相比，从头预

测方法不需要已知结构信息，直接从蛋白质序列预测其空间结构，因而在理论上是一种理想的方法，是一种理论科学与工程学相结合的方法，从蛋白质序列预测空间结构是蛋白质结构预测的最终目标。1973 年 Anfinsen 提出蛋白质的天然构象对应自由能最低时的结构这一热力学假设，是从头预测方法预测蛋白质空间结构的理论依据，然而这种方法一直受到两方面的困扰：首先，难以找到一个能严格区分蛋白质的天然构象和非天然构象的能量函数，使能量函数的全局极小点能对应于蛋白质的天然结构，目前已有的能量函数都采用经验势函数；其次，由于蛋白质系统相对分子量和柔性都很大，所以在蛋白质的势能面上存在着极多的局部极小点，缺少一种有效的全局优化算法也制约着这种方法的发展。

目前仍然不存在一个可信而又具有代表性的能量函数来预测出一个天然的折叠，也没有一个稳定的搜索方法能保证在对空间构象取样时，取样片段具有显著的天然性。对于预测折叠模拟的能量函数包括基于原子势能的一些分子力学软件包，如 CHARMM、AMBER、ECEPP 和平均势能的方法，还有基于多重序列比对的非物理的空间限制作用等。一些优化方法也是用于从头预测的，例如分子动力学模拟（molecular dynamics，MD）是模拟蛋白质分子以及蛋白质-底物复合物结构，可以根据蛋白质的功能结构模拟出一张关于蛋白质内部原子相互作用的热力学图谱及相关细节；蒙特卡洛模拟（Monte Carlo，MC）方法对分子间力应用不多，而主要应用于能量的比较；遗传算法（genetic algorithms）在蒙特卡洛模拟方法的基础上，对其预测时的收敛度加以改进，同时对取样及对方格法建立的模型都有所改进；还有运用扩散方程的方法。

4. SWISS-Model 简介

SWISS-Model（www. expaSy. org/swissmod/SWISS-MODEL. html）是日内瓦生物医学研究所（Geneva Biomedical Research Institute）建立的分子模建服务系统（molecular modeling server），采用一系列工作软件，从经典的布鲁克海文蛋白质结构数据库（Brookhaven PDB）中提取蛋白质查询序列的模拟结构信息，用具有蛋白质相似性的已知结构蛋白来建立未知结构蛋白的分子模型。目前 SWISS-Model 不仅能够产生一系列蛋白分子模型，而且与 SWISS-2DPAGE 等数据库建立了数据整合连接，扩大了服务系统的功能。SWISS-Model 模建服务系统界面见图 14-16。

SWISS-Model 提供以下服务。

（1）首选方式（first approach） 提交一未知结构 FASTA 格式序列（拟进行分子模建），服务系统将运用 FASTA、BLAST、SIM 和 CompAli 方法对晶体结构库序列进行一致性序列分析（sequence alignment），如果可行，ProMod 程序将建立分子的原子组成模型。

（2）优化模式（optimise mode） 将序列一致性序阵文件（alignment file）及一命令文件（command file）提交给 ProMod，可随意改变 SWISS-Model 推荐的序阵或减少模板结构数，取得更合适的效果。

（3）寡聚肽链模建（oligomer modelling）

（4）GPCR 模式（GPCR mode）

GPCR 是 G 蛋白偶联受体的 7 次跨膜结构模建。

SWISS-Model 整个分析步骤分为 5 步：①先决定序列是否能建模，SWISS-Model 将序列与晶体图像数据库（ExPdb）中数据进行比较，找出所有相似的同源序列；②检测序列同源性；③产生一个 ProMod Ⅱ 输入文件；④用 ProMod Ⅱ 建立三级结构模型；⑤用 Gromos96 算法修正，使模型的能量最小。用户可先选择“First Approach Mode”，呈现输入页面，输

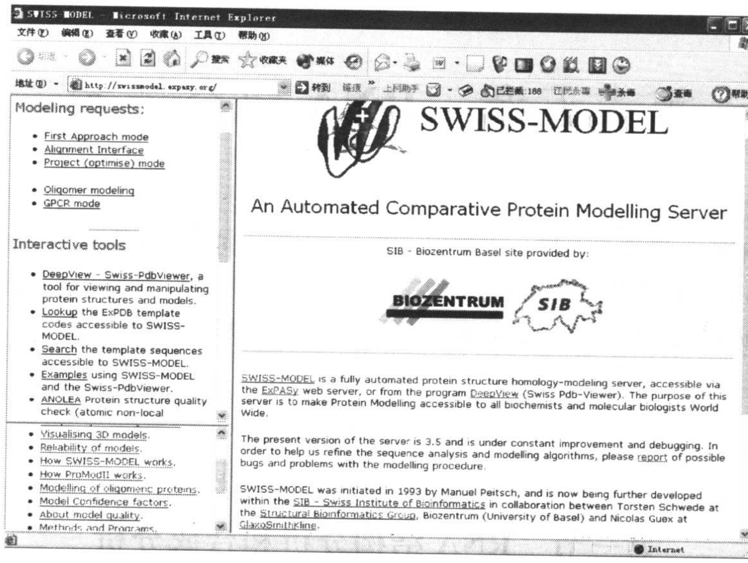


图 14-16　SWISS-Model 模建服务系统界面

入 E-mail 地址和查询序列，如果在数据库中有相当的同源序列，SWISS-Model 同意建模，会建立一个原子模型，并将结果以 E-mail 形式返回。然后可将这些结果再提交给 SWISS-Model 的“Optimize”模式，修正提出的结构模型。

预测结果根据用户提供的模板而定，即结果中会将所有模板及预测的三级结构都列出来。可以用 Swiss-PdbViewer（www. expasy. ch/spdbv/mainpage. htm）、Rasmol（www. bernstein-plus-sons. com/software/rasmol/）、VRML 插件（www. karmanaut. com/cosmo/player/）、Chime 插件（www. mdlchime. com/chime/）等来观看结果的三维结构图。

5. CPHmodels 简介

CPHmodels（www. CbS. dtu. dk/ServiceS/CPHmOdels/）是丹麦理工大学生物序列分析中心的预测服务器提供的三级结构预测。其方法是基于同源模建，与 SWISS-Model 相比，它简单易学，只需要提交序列，没有高级选项，如提交自定义的模板等。将待分析的序列提交后，可以随即得到结果，不会出现无搜索结果的情况。结果包括提交序列、模板序列、序列的双重比对、蛋白质位阻预测图，以及用 Javascript 显示的完整三级结构示意图。可以通过鼠标移动模型，同时可以下载 PDB 格式结构文件，注意这些文件是含有空位的三级结构。与 SWISS-Model 不同，它预测出的结果会根据目标序列与模板序列比对的空位，将空位通过断裂的方式表示在三维结构中，因此在阅读结果时应注意这些空位的位置。如果这些空位不是被预测蛋白的功能域，而是位于一些相对比较次要的位置，这样三级结构的整体特征就没有破坏，这时预测就比较成功。相反，如果这些空位是被预测蛋白的功能域，这时预测可能会有较大的误差。

（田　俊）

第十五章 RNA研究技术

RNA主要存在于细胞质中，RNA在遗传信息的传递方面起着承上启下的作用，其中rRNA、tRNA不能翻译成蛋白质，它们同mRNA一起共同参与蛋白质的合成。近来发现各种双链及单链小RNA参与了多种调控过程，具有广泛的作用，这些均成为RNA研究的重要内容，RNA研究技术也迅速发展起来，成为生物技术领域重要技术之一。

第一节 RNA结构分析和合成分析

一、用单链DNA探针和S1核酸酶分析mRNA

此方法利用寡核苷酸在激酶的作用下其5′端可以进行高效放射性标记的特点，将标记后的寡核苷酸与具有互补序列的特异性单链模板DNA杂交，并在*E. coli* DNA聚合酶Ⅰ的Klenow大片段的作用下使杂交体延伸；然后使用某种限制性核酸内切酶进行酶切，使探针具有明确的3′端结构，经碱性琼脂糖凝胶电泳分离得到探针。进行该方案前，最好能获得M13克隆，否则参照方案2使用含有目的序列的双链质粒来制备探针。方案3采用长寡核苷酸作为探针进行S1核酸酶分析，可用于对来源于特定启动子的mRNA水平做简单快速定量。

最后将得到的探针与RNA杂交来进行RNA的5′端作图，加入S1核酸酶消化探针中未与RNA杂交的部分，杂交体经变性聚丙烯酰胺凝胶电泳后，分析剩下的DNA片段的长度，此片段长度与探针的5′端距RNA的5′端间的长度相等，借此可判断核苷酸的转录起始位点。用足够数量的探针来进行杂交，就可以对样品中RNA的相对含量进行定量分析。

方案1 用M13模板进行mRNA的S1核酸酶分析

末端标记单链DNA探针的制备和S1核酸酶分析mRNA。

单链DNA探针的合成

① 用1×碱性凝胶缓冲液（50×，2.5mol/L NaCl，50 mmol/L EDTA pH8.0）制备1.2%低熔点琼脂糖凝胶，应使用较宽（每孔8mm）的梳子。胶凝固后，置于碱性电泳缓冲液（1.5mol/L NaCl，50mmol/L EDTA pH8.0）中浸泡过夜。

琼脂糖用碱性缓冲液煮沸时不会立即形成凝胶，必须浸泡过夜。也可以省去浸泡步骤但制备步骤较复杂，将琼脂糖加到水中煮沸，待冷却至50～60℃时加入碱性凝胶缓冲液使终浓度成1×，然后灌胶。

② 混合下述反应物制备激酶化的寡核苷酸，然后在37℃温育30min：

20μl [γ-^{32}P] ATP（10mCi/ml，6000Ci/mmol）

1μl 100μg/ml 寡核苷酸引物（100ng；最好20～30bp）

2.5μl 10×多聚核苷酸激酶缓冲液（700mmol/L Tris-HCl，pH7.5，100mmol/L $MgCl_2$，50mmol/L DTT，1mmol/L 氯化亚精胺，1mmol/L EDTA）

4U T4 多聚核苷酸激酶

使用的寡核苷酸必须能与目的 RNA 杂交，产生的探针在经 S1 核酸酶分析后能得到 50～250bp 长度的被保护片段。

③ 65 ℃加热 5min 灭活激酶。

④ 寡核苷酸与模板 DNA 退火反应。如果使用的是单链模板（如 M13 克隆），则在激酶化的寡核苷酸中加入以下混合物：55μl 水、18μg 模板 DNA 和 9μl 10×TM 缓冲液［每 100ml 含 1mol/L Tris-HCl（pH7.5）50ml，$MgSO_4 \cdot 7H_2O$ 2g，加水至 100ml，1.03×10^5 Pa 高压下灭菌 20min］。40℃杂交 15min。

⑤ 在上述杂交反应物中加入 9μl 4mmol/L dNTP 混合液（终浓度 400μmol/L）和 2μl *E. coli* Klenow 聚合酶，37℃温育 30min，使引物延伸合成全长探针。65℃加热灭活 5min，置于冰浴上。

⑥ 加入 40U 合适的限制性核酸内切酶和 10μl 相应的 10×限制性酶切缓冲液，37℃温育 45min，将探针切割成合适的长度。然后 65℃加热 5min 灭活内切酶。

⑦ 加 100μl 5mol/L 乙酸铵和 500μl 乙醇，－20℃沉淀过夜或在干冰/乙醇上沉淀 15min。

单链探针的分离

⑧ 碱性琼脂糖凝胶电泳：重悬探针于 25μl 碱性加样缓冲液（30mmol/L NaOH，1mmol/L EDTA，pH8.0，10% Ficoll，0.025% 溴酚蓝）中，然后小心上样，以最大 1.8V/cm 的电压进行电泳。

⑨ 从凝胶上分离探针：关掉电源，小心地使凝胶滑至一玻璃托盘中（在放射性屏障后操作），用塑料薄膜包裹凝胶，然后在暗室中将凝胶与 X 光片的一角对齐，曝光 3min 后洗片。

⑩ 将凝胶与 X 光片对比，上端较黑的带为探针，下端是未杂交的寡核苷酸，从探针所在位置切下凝胶块并尽量切掉多余的琼脂糖，使用小型探测仪监测，避免将探针随琼脂糖切掉。再将得到的长方形凝胶块从其底端开始切成细条，用小型探测仪检测每一细条，将计数较高的合并到 1 个或 2 个微量离心管中。

⑪ 65℃熔化收集到的凝胶条，用移液器确定体积后，加入等量的 TE 缓冲液，65℃加热 10min。

⑫ 加 1μl 10 mg/ml tRNA（将 tRNA 溶解于水中，终浓度 10mg/ml，酚抽提 2 次），用酚（注意不要用酚-氯仿）抽提两次，抽提时不要搅动界面白色物。氯仿会显著减少抽提产量。

⑬ 加 1/10 体积的 3mol/L 乙酸钠和 2 倍体积的乙醇沉淀，用 100μl 0.3mol/L 的乙酸钠重悬沉淀。

⑭ 取 1μl 用闪烁计数仪计数，确定每分钟计数值。

单链探针杂交和 S1 核酸酶消化

⑮ 在冰上将总量 5×10^4 cpm 计数的探针加至最多 50μg RNA 中，调整终体积为 100μl，乙酸钠浓度调为 0.3mol/L，再加 250μl 乙醇进行沉淀。用 70%乙醇洗涤沉淀一次，倒置于滤纸上干燥 30min。

⑯ 重悬沉淀于20μl S1 杂交液［80％甲酰胺，40mmol/L PIPES（pH6.4），400mmol/L NaCl，1mmol/L EDTA（pH8.0），分装成1ml/份，于－70℃储存］中，用移液器吸头上下吸 50 次以上并剧烈振荡，确保重悬完全。

⑰ 65℃变性 10min，30℃杂交过夜。

⑱ 第二日上午制备以下混合物（每一杂交反应体系）：

150μl 2×S1 核酸酶缓冲液［0.56mol/L NaCl，0.1mol/L 乙酸钠（pH4.5），9mmol/L $ZnSO_4$，过滤灭菌后 4℃储存］，3μl 2mg/ml 单链小牛胸腺 DNA，147μl 水，300U S1 核酸酶。

每一杂交反应体系中加 300μl 上述混合物，30℃温育 60min。

⑲ 每一反应体系中加 80μl S1 终止缓冲液［4mol/L 乙酸铵，20mmol/L EDTA（pH8.0），40μg/ml tRNA，4℃储存］，然后加 1ml 无水乙醇沉淀，用 70％乙醇洗涤沉淀一次，室温干燥 5min 。

⑳ 重悬沉淀于 3μl TE 缓冲液［10mmol/L Tris-HCl，1mmol/L EDTA（pH8.0）］中，加 4μl 甲酰胺加样缓冲液［0.2ml 0.5mol/L EDTA（pH8.0），10mg 溴酚蓝，10mg 二甲苯青，10ml 甲酰胺］，煮沸 3min 后置于冰上。

㉑ 取 3～5μl 样品在变性聚丙烯酰胺/尿素凝胶电泳，观察并分析预计大小的被保护带所占比例。

方案 2　以双链质粒为模板合成单链 DNA 探针

双链质粒也可用作制备单链 S1 探针的模板，但首先应按下述方法进行变性。

① 取 18μg DNA，加足量的 10×NaOH-EDTA 溶液［2mol/L NaOH，2mmol/L EDTA（pH8.0）］至 1×终浓度，室温下放置 5min。

② 加 1.5 倍体积的 4mol/L 乙酸铵（pH4.5）中和溶液，加 2.5 倍体积乙醇，－70℃沉淀 15min。

③ 离心收集沉淀，用 70％ 乙醇洗涤一次，重悬沉淀于 55μl 水中，余下按方案 1 中步骤④进行。

方案 3　用寡核苷酸探针和 S1 核酸酶定量分析 mRNA

在已知 RNA 结构的情况下，用此方案来测定 RNA 的量就很理想。除使用合成的 ^{32}P 标记 5′端的寡核苷酸（40～80mer）作为探针外，其余与方案 1 类似。

（1）设计寡核苷酸

① 对每一个 RNA 样品，寡核苷酸应至少包含 40 个与 RNA 编码链互补的碱基。寡核苷酸的 5′端与 RNA 互补尤其重要，而 5′端含 dG 或 dC 更好。

RNA：DNA 杂交体形成的速度和最佳结合状态及其稳定性与二者结合区段的长度有很大关系，以含 40～80 个互补碱基的寡核苷酸比较好，其 5′端为 dG 或 dC 可最大限度地降低 RNA：DNA 杂交体两端间的冲突。

② 如果实验目的是为了了解 RNA 的 5′端结构，则寡核苷酸的 3′端至少应延伸至 RNA 编码区以外 4 个碱基（如 RNA 起始位点上游）。如果要对不同 5′端的 RNA 水平进行定量，则设计的寡核苷酸至少应含有 40 个与样品 RNA 互补的碱基，这样可以最大限度减低 RNA：DNA 杂交体物理特性的差别。

40mer 以外的其他核苷酸可以使 RNA：DNA 杂交体形成的条带与未杂交的探针条带区别开来。如果不同的 RNA 样品其 5′端比较接近（20～30 个碱基以内），则可以使用同一条

探针，否则必须分别各自设计探针。

③ 用于检测对照 RNA 的探针，其 5′端结构并不重要，因此可以选择与 RNA 中间序列互补的寡核苷酸，但寡核苷酸的 3′端应含有至少 4 个额外的不与 RNA 互补的核苷酸，按互补的嘌呤和嘧啶进行选择。

在对照 RNA，选择其中间序列设计探针的最大优点在于形成的 RNA：DNA 杂交体在放射自显影片上形成的是单一条带，最大限度地减低了非互补核苷酸的不匹配现象，简化了 S1 消化和区别 RNA：DNA 杂交体与探针条带的步骤。

④ 寡核苷酸设计的要点是使从预计的转录物形成的 RNA：DNA 杂交体的大小足以通过凝胶电泳区别开来。在不同样品间能区别开的最小差别为 4bp。

(2) 制备杂交探针

① 准备以下 T4 多聚核苷酸激酶反应体系，终体积均 25μl。

2pmol dNTP

150μCi [γ-^{32}P] ATP (3000～7000Ci/mmol)

2.5μl 10×T4 多聚核苷酸激酶缓冲液

10U T4 多聚核苷酸激酶

使用尽可能高比放射活度的 [γ-^{32}P] ATP，以使敏感性达最大。无载体 [γ-^{32}P] ATP 的粗提物与其纯化物效果相当，而后者的比放射活度还低一些。如果寡核苷酸的来源有限或者待分析的样品量不多，则寡核苷酸的用量可以减低。换算公式：1pmol 寡核苷酸（长度 45mer)＝15ng 寡核苷酸。

② 37℃温育 30～60min，然后 75℃加热 10min 终止反应。

③ 加入 1μl 10mg/ml tRNA 、26μl 4mol/L 乙酸铵和 110μl 乙醇进行沉淀；用 26μl 水重悬样品，再加 26μl 4mol/L 乙酸铵和 110μl 乙醇，重复沉淀一次。

用 2mol/L 乙酸铵进行乙醇沉淀可以将寡核苷酸沉淀下来，未整合的 ATP 则不被沉淀。也可以用柱层析的方法分离未整合的 ATP 和标记的寡核苷酸，但寡核苷酸比较小时，必须使用如 Biogel P-2 类型的树脂。此方案可以得到足量探针，至少可进行 50～100 次杂交，但应在 6 周内用完（－20℃储存)，否则对结果的质量会有较大影响（当然，可以通过延长曝光时间来弥补放射性的衰减)。

(3) 杂交、S1 核酸酶处理和产物分析

杂交、S1 消化和产物分析可以完全按照方案 1 进行（从步骤⑮开始)，但必须注意杂交温度和时间，以确保杂交反应进行完全，或者按下述采用水性杂交体系。

① 每一待分析的 RNA 样品，准备以下 30μl 杂交反应体系。

20μl RNA（至少含 50ng RNA)

9μl 3×水性杂交液 (3mol/L NaCl，0.5mol/L HEPES pH7.5，1mmol/L EDTA pH 8.0)

1μl 探针混合物（每一种寡核苷酸 0.3ng，约 10^5cpm)

可以制备好足量用于所有杂交反应的杂交液和探针的预混合液。如果寡核苷酸的来源有限，可以大大缩减寡核苷酸的用量（但每种至少 0.1ng)，此时应设置对照以判断杂交反应是否完全。

② 75℃加热反应混合液 10min，然后 55℃温育过夜（或其他最佳杂交温度)。反应最好在温育箱或有盖的水浴槽中进行，以减少蒸发。

③ 短暂离心将液体甩至管底，然后置于 37℃。

④ 每一管中加 270μl 含 100～300U S1 核酸酶的 S1 核酸酶混合液（按方案 1 步骤⑱配制），37℃温育 30～60min，加 3μl 0.5mol/L EDTA、1μl 10mg/ml tRNA 和 0.7ml 乙醇终止反应。

⑤ 将混合液置于干冰上 10～15min 进行乙醇沉淀，70%乙醇洗涤一次，重悬于 10μl 0.1mol/L NaOH 中。

⑥ 取 3μl 重悬产物与 3μl 甲酰胺加样缓冲液混合，90℃保温 2min，进行聚丙烯酰胺凝胶电泳分析并曝光。

⑦ 使用密度测定仪扫描放射自显影片，检测目的 RNA 和对照 RNA 带的密度，这样可以用对照 RNA 量对目的 RNA 量进行校正。

方案 4　S1 核酸酶定量分析 mRNA 的调控

在进行精确定量时，以下几点非常重要：①探针必须过量；②杂交反应进行完全；③形成的 RNA：DNA 杂交体应很稳定；④S1 核酸酶反应恰当。在前面介绍的方案中，上述条件均应满足，但重要的是用适当的对照进行证实。首先，用不同量的 RNA 样品与等量探针杂交，确定形成的条带密度与加入的 RNA 量成比例，即成线性相关。其次，将杂交反应的样品按温育时间不同分成数份，每份均进行 S1 核酸酶消化，当最后的条带密度不随温育时间的延长而加深时，可认为杂交反应完全。在使用低浓度的标记核苷酸探针时，设置此对照尤为重要。第三，应改变杂交温度以确定最佳温度，温度应足够高使杂交反应高效进行，但又要足够低使 RNA：DNA 杂交体不会丢失（短杂交体的稳定性较长杂交体低）。第四，对标准的方案而言，还应采用不同的 S1 核酸酶用量，确保探针完全降解而不丢失目的信号。对一个 RNA 样品的分析，在已获得了上述参数后，在随后的分析中就不需重复进行。

S1 核酸酶分析可以应用以下方面：①用末端标记探针确定转录产物 3′端和 5′端结构；②对特定 RNA 的水平进行定量；③确定转录产物的方向；④在最初的真核转录产物中确定内含子的位置和大小。

S1 酶是一种单链核酸内切酶，对单链 DNA 和 RNA 都能进行消化。S1 分析的首要原则是 DNA 探针片段与细胞 RNA 杂交，然后再加入 S1 核酸酶消化未杂交的单链区，根据所用特异性探针的不同，可能是 5′突出端、3′突出端或内含子部分。双链 RNA：DNA 杂交体则能抵抗消化而保留下来，而标记的 DNA 片段则反映了与探针同源的杂交体中 RNA 的量和大小。

寡核苷酸的使用有几个优点：首先，探针制备迅速，用简单的 T4 多聚核苷酸激酶反应即可完成；其次，在激酶反应中用两种或以上等摩尔量的寡核苷酸，可以同时制备目的 RNA 和对照 RNA 的杂交探针，使用这样的探针混合物与 RNA 杂交可以测定中间序列相同的 RNA 的总量；第三，可以获得大量的寡核苷酸，保证探针量大大超过 RNA 而使杂交反应完全，而后者是进行精确定量的关键；第四，由于探针来源于合成的寡核苷酸，因而全部是单链结构，避免了因互补链存在引起的假象；最后，由于探针很短，即使 RNA 样品有轻度降解也能获得合理的结果。

S1 核酸酶分析最早由 Berk 和 Sharp 于 1977 年提出，开始时用 ^{32}P 标记高比放射活度的 DNA 与未标记 RNA 杂交来检测腺病毒早期转录产物，而用重叠的限制性酶切片段可以判断这些转录产物的大小和末端结构。此后，Weaver 和 Weissman 对原始方案进行了修改，他们采用末端标记的探针，可以同时检测转录的 RNA 链和通过 S1 降解后被探针保护的部分，从而可以直接测定距离。稍后，Favaloro 等又介绍了通过二维凝胶电泳快速分析内含子

的方法，以及通过一维S1凝胶经直接自显影或Northern blots来分析RNA的方法。

在进行S1核酸酶分析时，常常出现两个截然相反的问题：无信号或背景过高。没有信号产生的原因主要有3个：①探针的比放射活度没有达到要求，0.1μg的单链探针其要求比放射活度应达到每分钟计数 2×10^7（400base的单链探针比放射活度应达 10^7cpm/μg），如探针的比放射活度明显低于此，一种可能是激酶反应未能有效进行，原因可能是寡核苷酸在制备过程中残存的一些化合物抑制了激酶的活性，有时通过在寡核苷酸溶液中加入10mmol/L $MgSO_4$，然后加入5倍体积的乙醇进行沉淀可以除掉这些残余物；②RNA容易降解，可以将制备物进行琼脂糖凝胶电泳分析并用溴化乙锭染色了解有无RNA降解；③存在的目的RNA量非常少。如果是用转染的真核细胞进行S1分析，则目的RNA水平可能很低，可以通过oligo（dT）方法进行选择来增加其水平。也可能转染效率本身很低，应对转染方法进行改进。采用SP6探针可以增加敏感性10倍，从而可以检测到信号。

背景过高的原因之一是探针制备物中含有与探针同源的DNA，制备比单链模板短的探针（如400base或更少）以及从凝胶上切下探针时小心操作，避免未标记高分子量DNA的污染可以解决此问题。另一原因可能是S1核酸酶消化不完全，采用tRNA作为对照与探针进行杂交可以对此给出一些提示。除上述两种原因外，RNA样品广泛降解也有可能。

采用含甲酰胺的S1杂交液，杂交时RNA和探针混合物可能不会溶解，这样会产生很紧密的探针条带而信号很低。如果RNA样品完全干燥，则很难再被重悬，应让其沉淀保持一定湿度。

用0.1μg的寡核苷酸作为引物应得到每分钟计数 $(1\times10^7)\sim(2\times10^7)$/μg的单链探针，从低熔点琼脂糖凝胶分离DNA的效率大约是50%，因此估计合成的全长400base的探针其比放射活度约每分钟计数 $(1\times10^7)\sim(2\times10^7)$/μg，应使用足够放射活度的探针以保证检测到目的转录产物。随贮存时间的延长，探针的放射性会衰减而造成结果显示不清。使用变性胶的浓度根据检测目的带大小而定，可参见表15-1。

表15-1 变性胶浓度与目的带大小

聚丙烯酰胺尿素凝胶/%	目的带大小/base	聚丙烯酰胺尿素凝胶/%	目的带大小/base
4	>200	8	40～100
5	80～200	12	10～50

二、核糖核酸酶保护测定

1. 核糖核酸酶保护测定过程

将目的基因（即探针序列）克隆到噬菌体启动子下游多克隆位点上，制备具高特异性和高比放射活度的RNA杂交探针。先用限制酶酶切消化质粒，然后用噬菌体RNA聚合酶转录质粒DNA，合成高特异性和高比放射活度的RNA分子。对获得的RNA探针进行纯化，用DNA酶除去模板DNA，用酚-氯仿抽提除去蛋白质，再在乙酸铵存在的条件下经乙醇沉淀除去未整合的标记物；也可用凝胶电泳的方法纯化RNA探针。探针RNA与样品RNA进行杂交，杂交反应完成后用RNA酶除去未杂交的探针，剩下完整的探针RNA与样品RNA杂交体片段。经乙醇沉淀回收杂交体，再在序列胶上电泳进行分析，如果待检样品中含有目的mRNA，则可见到与探针长度一致的片段。

[α-^{32}P] CTP（10mCi/ml，400～800Ci/mmol 。有以水溶剂形式出售的商品，浓度一

般是10mCi/ml，使用的比活度应为400～800Ci/mmol，更高比活度产生的探针很不稳定。也可以用标记的GTP或UTP取代，但相应地要变更NTP混合液中另外3种未标记核苷酸的组成）。

制备探针

① 在一个高压灭菌的微量离心管中混合下述反应物。

5×转录缓冲液［T7、T3 RNA聚合酶：200mmol/L Tris-HCl（pH8.0），40mmol/L $MgCl_2$，10mmol/L亚精胺，250mmol/L NaCl。SP6 RNA聚合酶：200mmol/L Tris-HCl（pH8.0），40mmol/L $MgCl_2$，10mmol/L亚精胺。应使用由DEPC新处理的水新鲜制备］4μl，200mmol/L DTT 1μl，3NTP混合液（ATP、UTP和GTP各4mmol/L，0.5mmol/L EDTA，－20℃储存）2μl，［α-^{32}P］CTP（10mCi/ml，400～800Ci/mmol，有以水溶剂形式出售的商品，浓度一般是10mCi/ml，使用的比活度应为400～800Ci/mmol，更高比活度产生的探针很不稳定。也可以用标记的GTP或UTP取代，但相应地要变更NTP混合液中另外3种未标记核苷酸的组成）10μl，RNA酶抑制剂1μl（20～40U），0.5mg/ml模板DNA1μl（终浓度25μg/ml），噬菌体RNA聚合酶1μl（5～10U）。

加样时应在模板DNA加入前先加入前4种成分，以防止转录缓冲液中含有的亚精胺沉淀DNA。

② 如用SP6 RNA聚合酶，在40℃温育；如用T7和T3 RNA聚合酶，则在37℃温育，时间30～60min。

标记的三磷酸核苷会迅速消耗，因此不必要延长温育时间，增加酶用量也没有更多帮助。实际上，存在过多的酶会导致模板的随机转录，而这些转录物很可能与探针互补结合，使杂交时的背景增高。

③ 加5μg或10U无RNA酶的DNA酶，一般用5000U/ml或2.5mg/ml的储备液2μl，37℃温育15min。这一步消化除去模板DNA，对杂交的成功非常关键。

④ 加2μl 10mg/ml tRNA作为载体，然后加水使终体积为50μl。

⑤ 用酚-氯仿-异戊醇抽提。

⑥ 回收水相并加入200μl 2.5mol/L乙酸铵和750μl无水乙醇，冰浴15min，4℃离心15min沉淀RNA。

如果准备用胶纯化RNA探针，则下一步接方案2。

⑦ 用50μl水重溶沉淀，加200μl 2.5mol/L乙酸铵和750μl无水乙醇，同样按步骤⑥沉淀。

⑧ 重复步骤⑦。

三次沉淀可去除几乎所有未结合的标记物，也可以采用其他分离RNA和游离三磷酸核苷的方法（如离心、柱层析法），但要十分小心，避免RNA酶的污染。

⑨ 用75%乙醇-25%乙酸钠洗涤沉淀。

⑩ 干燥后溶于100μl杂交缓冲液［5×储存液：200mmol/L PIPES（pH6.4），2mol/L NaCl，5mmol/L EDTA。使用时，应用冻存的5×储存液1份和新经混合树脂去离子处理的甲酰胺4份（1∶4）配置新鲜杂交工作液，也可将工作液分装成小份在－70℃保存］中。

⑪ 取与1μl进行液体闪烁计数。

探针的比活度约有每分钟计数10^9/μg，由于放射性衰减，其活度迅速下降，因此在制备当天使用效果最佳，如在4℃储存，也应在几天内使用。一周后，探针大部分降解。

探针RNA与模板RNA杂交

⑫ 用无水乙醇沉淀RNA样品，如果是储存于水中的RNA，则需冻干。

对细胞总RNA或胞质RNA，通常10μg就携带有大部分信息，如目的RNA丰度较高，用量还可以少一些。另外，应以含tRNA的样品作为对照，一方面显示杂交背景强度，一方面检验RNA酶消化情况，一般杂交反应不会产生探针保护现象。

⑬ 重新溶解于含每分钟计数 5×10^5 个RNA探针的30μl杂交缓冲液中。

注意确保RNA沉淀完全溶解，可用吸管重复抽吸。杂交缓冲液中如有一个以上的探针，可对单个样品的多种mRNA进行分析。随缓冲液中探针含量的增加，杂交背景也随之增大，因此不必要过多增加探针用量。

⑭ 85℃加热5min使RNA变性。

⑮ 迅速转至设定的理想杂交温度，温育过夜（8h以上）。

由于在S1核酸酶分析中RNA探针不会像双链DNA那样自我退火，因此杂交温度的要求不很严格，但对不同的探针仍有一最佳杂交温度，原因可能在于探针二级结构的形成。建议温度范围为30～60℃，可先试45℃杂交。对某些RNA，杂交可以在4～6h内完成。

RNA酶消化

⑯ 每一杂交反应体系中加入350μl含40μg/ml RNA酶A和2μg/ml RNA酶T1的RNA酶消化缓冲液[10mmol/L Tris-HCl（pH7.2），300mmol/L NaCl，5mmol/L EDTA，加1/50体积的50×RNA酶混合液（2mg/ml RNA酶A，0.1mg/ml RNA酶T1）。RNA酶消化缓冲液应在室温下储存，用时加入需要量的RNA酶冻存液，配制此溶液一定要使用一次性试管]，30℃温育30～60min。此次的温育条件对最终杂交信号的强度影响不大，但如果怀疑此步有问题，则应该从15～37℃温育范围进行检测，如果发现RNA复合体双链结构内部有断裂，则去掉消化缓冲液中的RNA酶A。

⑰ 加10μl 20%SDS和2.5μl 20mg/ml蛋白酶K，37℃温育15min。

⑱ 用400μl酚-氯仿-异戊醇抽提一次，水相移入含有1μl 10mg/ml酵母tRNA的微量离心管中。

⑲ 加1ml乙醇，沉淀。

⑳ 干燥沉淀后，重溶于3～5μl RNA加样缓冲液[80%（体积分数）甲酰胺，1mmol/L EDTA（pH8.0），0.1%溴酚蓝，0.1%二甲苯青，不要使用Maxam-Gilbert加样缓冲液或任何含有NaOH的缓冲液]中。

㉑ 85℃温育3min变性。

㉒ 在变性的聚丙烯酰胺-尿素凝胶（测序胶）中电泳分析结果。

在这种凝胶中RNA的电泳迁移速率较同样大小的DNA慢，如果使用DNA分子量标准来估计RNA片段的长度，则RNA的实际长度要较相当的DNA标准带小5%～10%。例如，如果RNA片段的迁移速率与100bp大小的DNA相当，则RNA片段的实际长度约90～95 bp。

2. 凝胶纯化RNA探针

某些条件下如需要全长的RNA探针时，可以使用本方案进行纯化。

① 在第一次乙醇沉淀后，晾干，重溶于10μl RNA加样缓冲液。关键在于RNA完全干燥，然后再充分溶解。

② 85℃加热5min使RNA变性。

③ 加样至 TBE 缓冲液制备的 6% 聚丙烯酰胺凝胶（丙烯酰胺与亚甲双丙烯酰胺比例为 29：1），凝胶一般厚 0.4mm、长 14cm（也可以使用测序胶长度），300V 电压下电泳（如胶更长，则电压也应增高），直到指示剂溴酚蓝泳动到凝胶的 1/2～2/3 位置。

使用非变性胶在于其泳动速度快，但一定要注意在加样前使 RNA 完全变性，也可以使用含尿素的变性胶。

④ 拆开凝胶夹层，让凝胶保留在一面玻板上，用放射性墨水作好标记，再用塑料薄膜包裹好，感光胶片曝光 30s，以胶片作对照，切下全长的 RNA 带。

RNA 带在凝胶上所占实际位置较胶片上显示的带要细很多，因此不要切下很多凝胶，切下的凝胶越小，杂交越干净。

⑤ 加 400μl 洗脱缓冲液（2mol/L 乙酸铵，1% SDS，25μg/ml tRNA），37℃摇动 2～4h 洗脱 RNA。

洗脱缓冲液在室温下常出现沉淀，因此在使用前应稍加热。较小的探针（<200base）洗脱 90min；较长的核酸应洗脱更长时间，但一般 2h 应足够。

⑥ 将洗脱液移到一新微量离心管中，加 1ml 无水乙醇。

⑦ 冰浴 15min，离心 15min。

用探测仪检测洗脱液和切下的凝胶，前者的计数应高很多，且应被乙醇沉淀下来。

⑧ 重溶 RNA 于 50μl 杂交缓冲液中，取 1μl 进行液体闪烁计数。

产量会比未用胶纯化前低一些，但所得探针足以进行 50 次杂交，而背景会减低很多。

3. 模板 DNA 的合成

模板 DNA 是通过将目的序列插入到含噬菌体启动子的质粒载体中来进行制备，含有噬菌体启动子的载体有商业化产品出售。

用限制酶在目的探针序列的下游处切断，这样可以产生均一长度的转录产物。所选用的限制酶可以在质粒上有多个酶切位点，但不能在噬菌体启动子、探针序列和两者间的载体 DNA 内有切点。选用的限制酶以能产生 5′突出端者为最佳，如采用 3′突出端产生的限制酶，由于 3′突出端也可作为 RNA 聚合酶的起点，可导致合成与探针互补的 RNA 链。RNA 的消化要彻底，但不要过度。用酚-氯仿抽提酶切后的 DNA，乙醇沉淀后，重新溶于无 RNA 酶的 TE 缓冲液中（0.5mg/ml），一般少至 100ng 的模板 DNA 就可以产生相当量的探针。

噬菌体 RNA 聚合酶具有的特性使其非常适合用于制备高比活度的探针。首先，这种酶相对比较稳定且易于纯化，而且其基因已经整合到噬菌体基因组中，可以更直接地在大肠杆菌中进行克隆和表达，因此减轻了纯化难度并提高了效益。其次，其合成 RNA 的速度非常快，可达 200～300base/min，比大肠杆菌 RNA 聚合酶和 DNA 聚合酶的速度约快 10 倍，因此这种酶可用于合成大量的探针。第三，此酶识别的启动子序列很长，不可能偶然地与其他 DNA 结合，特异性非常强。因此，用此酶合成的探针与原序列有很好的同源性，并且纯化简单。

三、引物延伸法

本方案用于检测 RNA 的量和 5′端长度，用 30～40mer 末端标记的寡核苷酸作为引物与待检 RNA 杂交，而后加入脱氧核苷酸延伸此引物，合成与模板 RNA 互补的 cDNA，然后对此 DNA 进行序列胶电泳分。延伸引物的长度反映了引物与 RNA 5′端的距离，而延伸

引物的量则反映了待检测信息的水平。

寡核苷酸的标记

① 混合以下试剂，37℃温育30min，制备标记的寡核苷酸。

20μl [γ-^{32}P] ATP (10mCi/ml，共200μCi)

1μl 100μg/ml寡核苷酸引物（100ng，特异性针对目的RNA序列）

2.5μl 10×多聚核苷酸激酶缓冲液

4U T4多聚核苷酸激酶

采用30～40mer的寡核苷酸引物，合成的延伸产物长度应为100base左右。如果使用更短的寡核苷酸，应适当放宽杂交条件，如降低杂交温度或在水溶液中进行杂交。引物延伸存在的一个主要问题是逆转录酶会在RNA的某一序列处终止合成，因此预计合成的产物不能太长，这样危险性就低一些。

② 65℃加热5min灭活激酶，加25μl 4mol/L乙酸铵和250μl乙醇，在干冰-乙醇上沉淀30min或－20℃沉淀过夜，然后离心15min得到标记的寡核苷酸沉淀。

③ 用25μl水溶解寡核苷酸沉淀，再加25μl 4mol/L乙酸铵，重复上述沉淀过程2次。3次沉淀的目的在于从标记的寡核苷酸中除掉[γ-^{32}P] ATP。

④ 用100μl 0.3mol/L乙酸钠重溶寡核苷酸，取1μl进行计数。

杂交

⑤ 将计数为5×10^4的标记寡核苷酸和最多50μg RNA混合，调节溶液的盐浓度至终浓度为0.3mol/L乙酸钠，再加2.5倍体积的乙醇进行沉淀。倒转离心管置于滤纸上干燥30min，重悬于30μl 1×水性杂交液或30μl S1杂交液中，用移液器吸头上下抽吸液体至少50次并充分振摇以确保RNA完全溶解。然后在30℃温育过夜。

RNA完全干燥后再进行重悬非常困难，因此不要真空干燥RNA沉淀。杂交条件应根据引物的长度和所含GC量来定，较长的引物（>30mer）应用甲酰胺杂交液[80%去离子甲酰胺，40mmol/L PIPES (pH 6.4)，400mmol/L NaCl，1mmol/L EDTA (pH 8.0)，分装成1ml，于－70℃储存]，较短的引物则用水性杂交液[1mol/L NaCl，0.33mol/L HEPES (pH7.5)，033mmol/L EDTA (pH8.0)]。

⑥ 在杂交反应体系中依次加入170μl 0.3mol/L乙酸钠和500μl乙醇，沉淀RNA和寡核苷酸。用75%乙醇-25% 0.1mol/L乙酸钠(pH5.2)洗涤沉淀，倒置于滤纸上干燥。

引物延伸

⑦ 在冰上新鲜制备RT混合液（按每一反应计算）。

3.5μl 4mmol/L dNTPs

2.5μl 10×RT缓冲液

1.25μl RNAsin

18μl 水

每一样品中加入25μl RT混合液，充分重悬，通过监测放射比活度可以了解RNA和寡核苷酸的重悬情况。

⑧ 每一反应体系中加2μl (40U) AMV逆转录酶，42℃温育90min。温度低于42℃时，逆转录反应更容易终止。

⑨ 每一反应体系中再加1μl 0.5mol/L EDTA，然后加1μl 1mg/ml RNA酶A，37℃温育30min。当待分析的RNA样品超过10μg时需用RNA酶进行消化，目的是防止在用序列

胶分析样品时出现误差。

⑩ 加 100μl 2.5mol/L 乙酸铵，用 125μl 酚-氯仿-异戊醇抽提，将上层水相转至一干净管中，再加 300μl 乙醇沉淀 RNA。加 500μl 70%乙醇洗涤 RNA 沉淀，并干燥沉淀。

⑪ 重悬沉淀于 3μl TE 缓冲液中，加 4μl 甲酰胺加样缓冲液［0.2ml 0.5mol/L EDTA (pH8.0)，10mg 溴酚蓝，10mg 二甲苯青，10ml 甲酰胺］，煮沸 3min 后置于冰上，取 3～4μl 在适当比例的序列胶上进行分析。

本方法通常用于同时测定 RNA 的量和大小。逆转录酶的发现为此奠定了基础，此酶可以转录 RNA 为 DNA，因此单链 DNA 与 RNA 杂交后在此酶的作用下就可以沿 RNA 延伸直到达 RNA 的末端。

引物延伸法的主要困难在于逆转录酶遇到 RNA 的某些区域时（二级结构较复杂的区域）会暂停或终止作用，在进行凝胶电泳时就出现一些低分子量的条带，导致对结果解释不清并减少合成的全长延伸引物的量。延伸模板链中 GC 含量和发夹结构较多处可形成这样的终止或暂停位点。本方法所用核苷酸的浓度和温育温度以尽量防止此问题出现。使用的引物起始位点应距待分析信息末端 100base 以内，使逆转录酶合成距离不长也有助于减少这一问题的发生。由此产生的另一个问题是如果待分析的 RNA5′端正好是富含 GC 的区域，实验者就必须判断其结果究竟是真实的 5′端还是由于逆转录酶不能通过该区域而产生的假象，这时可联合使用 5′端标记的探针进行 S1 分析法来对 5′端定位。

四、RNA run-on 分析

在细胞核提取液中加入放射性同位素标记的 NTP，使其掺入到正在转录的 mRNA 分子中，通过核酸分子杂交分析，即可同时鉴定多种不同基因是否转录及其转录活性的高低。

(1) 细胞核的分离

① 培养的细胞，10ml 冰 PBS 洗涤 3 次，每培养皿加入 1ml PBS，用细胞刮棒将细胞刮入离心管中，1000r/min，4℃，5min 离心，弃上清。

② 将细胞悬浮于 10ml 10mmol/L Tris-Cl，pH 7.4，10 mmol/L NaCl，3mmol/L $MgCl_2$，0.5% NP-40 中，细胞浓度为 1×10^7 个/ml 左右。

③ 冰浴 10min，4℃，3000r/min，5min，弃上清。

④ 细胞重悬浮于 10ml 10mmol/L Tris-Cl，pH 7.4，10mmol/L NaCl，3mmol/L $MgCl_2$，0.5% NP-40 中，4℃，3000r/min，5min，弃上清。

⑤ 沉淀悬浮于 100μl 50mmol/L Tris-Cl，pH 8.3，5mmol/L $MgCl_2$，0.1 mmol/L EDTA，40%甘油中；−80℃保存。

(2) 标记 RNA 的制备

① 100μl 上述细胞核提取物，加入 100μl 反应液：10 mmol/L Tris-Cl，pH 8.0，5mmol/L $MgCl_2$，300mmol/L KCl，0.5mmol/L ATP，0.5 mmol/L CTP，0.5mmol/L GTP，3.7mmol/L Bq［α-^{32}P］UTP (111 TBq/mmol/L)，30℃，30min。

② 加入 4μl 1mg/ml DNase I，30℃，10min。

③ 加入 200μl 终止液 (20mmol/L Tris-Cl，pH 7.4，10 mmol/L EDTA，2% SDS，200μg/ml 蛋白酶 K)。

④ 50℃，1h。

⑤ 酚-氯仿抽提，将上清液转移到 10ml 离心管中；加入 50μg 酵母 tRNA，及 5ml 预冷

的 5% TCA-50mmol/L $Na_4P_2O_7$。

⑥ 冰浴 30min。

⑦ 用直径 2.5cm 的硝酸纤维素膜过滤，收集不溶成分，用预冷的 5% TCA-50mmol/L $Na_4P_2O_7$ 洗涤试管 1 次，并过滤。

⑧ 用预冷的 5% TCA-50mmol/L $Na_4P_2O_7$ 洗涤滤膜 3 次。

⑨ 将硝酸纤维素膜转移至闪烁瓶（或其他容器中），加入 0.9ml 保温液（20mmol/L HEPES，pH 7.5，5 mmol/L $MgCl_2$，1 mmol/L $CaCl_2$，25μg/ml DNase I）。

⑩ 37℃，30min。

⑪ 加入 30μl 0.5mmol/L EDTA（终浓度为 15mmol/L），100μl 10% SDS（终浓度为 1%）。

⑫ 65℃，10min。

⑬ 将反应后的溶液转移到 10ml 离心管中，再向装有滤膜的容器中加入 0.5ml 10mmol/L Tris-Cl（pH 7.5），1% SDS，5 mmol/L EDTA。

⑭ 65℃，10min。

⑮ 将上述两溶液合在一起，约 1.5ml，加入 40μg 蛋白酶 K（终浓度为 25μg/ml）。

⑯ 37℃，30min。

⑰ 加入 5mol/L NaCl，使终浓度为 0.1mol/L。

⑱ 分装至 4 个离心管中，加入 2.5 倍体积的冰乙醇，−80℃，10min。

⑲ 10000r/min，离心 10min，或 3000r/min，离心 30min，弃上清。

⑳ 风干后，溶于 50μl 10mmol/L Tris-Cl（pH 7.5)-2mmol/L EDTA 溶液中，即为标记的 RNA 探针。

（3）杂交测定　标记的 RNA 制成后，即可与特定的靶 RNA 或靶 DNA 进行杂交测定，具体方法可参见核酸杂交的相关章节。

第二节　RNA 组学技术

传统观念认为：三类最重要的生物高分子化合物中，DNA 携带遗传信息，蛋白质是生物功能分子，而 RNA 在这二者间起传递遗传信息功能（即参与蛋白质的生物合成）。基因组研究最终可提供全套的 DNA 序列，从中可检出所有的蛋白质基因。由于每个细胞都有全套遗传信息，但不同生长发育阶段、不同组织或不同生理病理情况下的蛋白质谱是不同的。20 世纪 80 年代初，T. Cech 发现 RNA 也可成为生物催化剂，他称之为核酶（ribozyme）。在酶学领域，核酶的发现打破了多年来“酶的化学本质就是蛋白”的传统观念。在 RNA 领域这一发现对传统观念的冲击更大。它使人们认识到，RNA 的生物功能远非“传递遗传信息”那么简单。1995 年提出蛋白质组学，以研究蛋白质在不同时空情况下的表达及其生物学意义。由于 DNA 与蛋白质的不完全对应性，1997 年又有人提出了转录组学（transcriptomics），以研究蛋白基因转录的时空关系及其生物学意义。但如上所述，RNA 的生物功能远超出了遗传信息传递中介的范围，中外科学家都注意到了以 RNA 的时空表达情况及其生

物学意义为研究对象的 RNA 组问题。国外在 2000 年底提出了 RNA 组学（RNomics）的概念。如果说基因组学研究正全力构筑生命科学基石的话，那么 RNA 组学研究和蛋白质组学、生物信息学等都是它的不可缺少的同盟军。

一、反义核酸技术

反义核酸技术是继基因克隆和重组技术后分子生物学领域兴起的一种全新的技术。该技术由于其核苷酸具有与 DNA 意义链相同的序列且可以选择性地抑制特定基因的表达而得名。广义的反义核酸技术包括：① 反义寡核苷酸技术（antisense oligo nucleotides, ASOD），即直接应用一段人工合成的寡聚核苷酸，通过碱基配对与细胞内核酸结合，特异地调节基因表达；②反义 RNA 技术（antisenseRNA），是利用基因重组技术，构建表达载体，使其离体或在体内表达出反义的 RNA。

目前，在基因功能的研究和基因治疗中，人们越来越广泛地应用反义核酸技术来抑制目的基因的表达，该技术在分子生物学和药物化学中有着巨大的发展潜力。

由于 ASOD 的易降解性，在实验中通常要对其进行修饰，硫代反义寡核苷酸作为第一代反义寡核苷酸的代表，由于在细胞毒性、细胞吸收率等方面存在众多问题，Agrawal 等设计了第二代混合骨架的反义寡核苷酸（mixed-backbone oligo nucleotides, MBO），MBO 含有至少两种不同的化学修饰，在结构上可以是任意两种或多种修饰的不同排列组合。MBO 不仅保持了核酸酶抗性，而且具有很好的靶序列杂交特异性和相对较低的毒副作用。近年来以 2-氨基乙基甘氨酸为基本单元的多肽核酸（peptide nucleic acid, PNA）由于可与互补的 RNA 或 DNA 形成稳定的 PNA/DNA（RNA）杂合链并且具有较强的蛋白酶、核酸酶抗性而预示了其作为反义抑制剂的用途，被称为第三代反义核酸，但细胞对 PNA 的吸收效率较低的问题一直没有得到很好的解决。并且，由于寡核苷酸的半衰期较短，有必要反复应用 ASOD。事实上，以上三代寡核苷酸较高的制备成本也是妨碍其广泛应用的因素之一。同时，在反义寡核苷酸的设计上还较盲目，要设计出高效的反义寡核苷酸并非易事。相对而言，包含反义 RNA 的重组质粒则相当于具有生物兼容性和生物降解性的长期释放装置。

对反义寡核苷酸技术而言，人们已在药代动力学、药效学和毒理学研究方面积累了一些实践经验，但其作用机制、特异性核苷酸设计、化学修饰和剂型设计、靶组织的有效给药方法、有效剂量与安全性问题尚需进行深入研究。而且，规模合成修饰的寡核苷酸的价格应进一步降低。反义 RNA 技术由于其成本较低及作用的长效性等反义寡核苷酸技术无法比拟的优势正逐渐受到重视。下面将对反义 RNA 技术作一简要介绍。

反义 RNA 是首先在细菌中被发现的一段很短的 RNA 序列，本身不被翻译，但是可以通过碱基互补与其他可翻译的 RNA 结合，从而阻止基因的正常表达。因此，反义 RNA 被认为是一种基因表达的抑制因子，它在复制水平、转录水平和翻译水平上都能对相关基因的表达进行调控。反义 RNA 技术作为反义核酸技术之一，已在越来越广泛的领域得到应用。反义 RNA 技术作为一种有力的工具，在基因功能研究方面，特别是在后基因组研究中将做出巨大贡献。

（一）反义 RNA 作用机制

与 ASOD 方法相比，设计反义 RNA 时一般不需要目的基因的详细的调控知识，人们普遍应用由翻译起始位点延伸 0.5～ 3.0kb 的片段作为重组质粒的（逆向）插入序列。而全长

的cDNA序列一般认为没有必要，尤其当目的基因较长时。另外，互补于3′端非翻译区的序列也往往具有很好效果。在真核生物中，一般认为对应5′非编码区的反义RNA要优于针对编码区的反义RNA。包含反义核酸的重组质粒在细胞内转录出反义mRNA，发挥抑制基因表达的作用。目前，具体的作用机制尚不甚清楚，人们推测可能有以下几条。

（1）在复制水平上 与引物RNA互补结合，抑制DNA复制，从而控制DNA复制效率。

（2）在转录后水平上 与mRNA 5′端互补，阻碍帽子结构形成，阻碍前mRNA剪接，阻碍于polyA形成位点，阻碍mRNA成熟及向胞浆转运。

（3）在翻译水平上 一般认为是较重要的作用方式——互补于AUG区域，阻止核糖体的结合；互补于编码区，则阻止核糖体的移动，或直接降解mRNA。这与人们普遍认为的ASOD的机制相似。反义RNA的重组体内含有启动子及终止子，当转染细胞后，重组体能自动表达反义RNA。重组体自身可能整合到宿主的基因组DNA中，或作为“附加体”长期存在。

（二）反义RNA技术方法

应用反义RNA方法大体经过以下三个步骤。

1. 反义表达载体的构建

一般应用RT-PCR技术克隆出部分目的cDNA，并在两端加上限制性酶切位点，倒向插入真核表达载体中，载体一般要含有下游的polyA尾以保证转录产物的稳定。为确定插入方向可应用限制性内切酶酶切或直接测序。

下面是一些常用的真核表达载体。

应用较广的是含SV40（simianvirus 40）、RSV（Rous sarcoma virus）或CMV（cytomegalo-virus）的启动子/增强子的质粒载体。尤其CMV的早期启动子/增强子元件可有效地应用于各种人及鼠源的肿瘤细胞系。含EBV（Epstein-Barr virus）复制起点的载体，可编码核抗原（EBV nuclear antigen），可以以“附加体”形式在细胞中复制，避免了整合人染色体基因组所带来的位置效应。在应用反义RNA技术研究原癌基因在细胞生长、分化及凋亡的调控机制时，人们发现反义RNA对一些原癌基因表达抑制的同时抑制了细胞的生长，从而不利于转化子的筛选。一种解决办法是选择可诱导型的表达载体，使得筛选后在诱导物的诱导作用下表达出反义RNA并发挥反义作用。1982年，Karin等构建含金属硫蛋白2型（MT2）启动子的可诱导型载体，在适当诱导条件下（如Zn^{2+}、Cd^{2+}等）才能启动下游基因的转录。应用此类载体进行转染已取得良好效果。构建此类载体时要注意的是诱导物对生物体可能产生的生物学效应，尤其是毒副作用。近年来出现了一类高效且诱导剂生物效应较小的双载体系统，如pVgRXR/pIND（Invitrogen）、pTet/pTRE（Clontech），其原理是第一种质粒能够表达能与诱导剂结合的蛋白，细胞在稳定转染了第一种质粒后，再转染第二种含目的基因片段或反义片段的质粒，而第二种质粒的表达受控于能被第一种质粒所表达的蛋白所激活的转录启动子。PTet/pTRE载体系统已成功地应用于细胞培养及转基因小鼠。但对于某些细胞系，四环素结合蛋白的高表达对其有一定的毒副作用。

2. 基因的导入

磷酸钙法、电转法、基因枪是体外实验中DNA转染真核细胞的常用方法。脂质体不仅常用于体外实验中载体DNA对细胞系的转染，也用于体内实验中的转染过程。脂质体能提高分裂细胞和非分裂细胞对核酸的摄取，并保护其不被降解；优点是操作简单、细胞吸收较好；缺点是包埋的DNA量不确定、体内实验中发现在特定组织如肝脏吸收的倾向性，及对某些组织如脑有一定毒性。在此基础上人们逐渐开发出具有一定良好特性的转染试剂如pH

敏感脂质体。各类受体介导的转染试剂往往使转染更接近自然的生理过程（如基于转铁蛋白受体介导的核酸转染试剂），不仅可能降低毒副作用还可因受体的特异分布而具有转染的组织或细胞特异性。另一种转化策略是以病毒作为载体，同时利用了病毒的转染特性。RAAV-AS（AAV，adeno-associated virus，腺病毒辅助性病毒）以其安全、高效而广泛应用于反义 RNA 技术及基因治疗。它几乎可以转染所有哺乳动物，但对其他动物如家禽，目前尚未建立起其 AAV 系；并且，由于它的转染是细胞膜上受体介导的，哺乳动物的 AAV 不能用于家禽的转染。逆转录病毒在快速分裂的细胞中转染效率也相当高，这使它常被作为载体应用于肿瘤的治疗中，它可应用于哺乳动物及禽类；其不足是它只在快速分裂的细胞中起作用，这使它的应用范围受到一定限制。一般认为体内的转染需要借助基因导入系统（gene deliver system），Wolff 和 Malone 等于 1990 年用表达 β-gal 及荧光酶基因的质粒分别对小鼠肌肉进行直接体内注射，在注射区域的肌细胞检测到该基因的表达。直接注射法仅对骨骼肌细胞有效，人们推测可能与骨骼肌的多细胞核、存在胞质环流等结构特性有关。为检验转染效率，人们往往选用 β-gal 或 pEGFP 等报告基因。

3. 检验、结果分析

在以反义 RNA 表达载体进行细胞系的转染或体内实验时，应设置空载体转染的对照实验，以排除外源载体 DNA 的导入带来的生物学或免疫学效应。体外实验中基因产物的活性、细胞的增殖、分化等生物学参数，体内实验中动物本身的一些生物学表型的改变均是不可忽略的重要指标。在此基础上分子生物学水平的分析主要有以下几个方面：除了以正义探针检验反义 RNA 的表达情况外，我们更关心的是靶基因的表达情况。应用进行中的核转录分析（nuclear run-on transcription assay）能够较可靠地测定基因转录活性。对 mRNA 水平的检验方法除了 Northern 印迹外，RNA 酶保护实验（RNase protection assay）灵敏度更高，并且可以进行较准确的定量分析。由于目前反义 RNA 抑制靶基因表达的原理尚在研究之中，而反义 RNA 抑制靶基因表达的过程中并不一定伴随着 mRNA 水平的降低，所以检测反义抑制效果的较为客观的标准是相应蛋白水平的变化。Western 印迹、ELISA、流式细胞分析法（flow cytometry）均是常用的分析技术。免疫组化技术、免疫荧光技术由于不能很好地定量而一般避免使用。

（三）反义 RNA 技术应用

反义 RNA 技术研究基因表达的高度专一性的反义抑制剂。在基础研究方面，反义 RNA 技术不仅用于基因分析，还可用来制备 C-缩短的多肽，以研究蛋白质结构活性关系而无需使基因变异；研究细胞信号传导、细胞分化及癌变的发生、发展过程等。在应用研究方面，反义 RNA 药物可以特异性地抑制癌基因的异常表达或抑制肿瘤细胞特异蛋白质的表达，诱导肿瘤细胞凋亡；在抗病原微生物感染方面，针对特异基因的反义核酸可较好地阻止其对人体的侵害；反义核酸技术用于抗癌、抗病毒的基因疗法方面已有大量报道。

二、RNA 干扰（RNAi）技术

RNA 干扰（RNA interference，RNAi）是正常生物体内抑制特定基因表达的一种现象，它是指当细胞中导入与内源性 mRNA 编码区同源的双链 RNA（double stranded RNA，dsRNA）时，该 mRNA 发生降解而导致基因表达沉默的现象。这种现象发生在转录后水平，又称为转录后基因沉默（post-transcriptional gene silencing，PTGS）。RNAi 发生于除原核生物以外的所有真核生物细胞内。需要说明的是，由于 dsRNA 抑制基因表达具有潜在

高效性，任何导致正常机体 dsRNA 形成的情况都会引起不需要的相应基因沉寂。所以，正常机体内各种基因有效表达有一套严密防止 dsRNA 形成的机制。

外源 dsRNA 进入细胞后产生的小分子干扰 RNA（small interfering RNA，siRNA）的反义链和多种核酸酶形成了沉默复合物（RNA-induced silencing complex，RISC），RISC 具有结合和切割 mRNA 的作用而介导 RNA 干扰的过程。生物体内的 DNA（脱氧核糖核酸）通常是双链的，而 RNA 则大多是单链的，基因是由 DNA 通过转录形成信使 RNA，随后翻译形成蛋白质才能起作用。小 RNA（microRNA、miRNA 或 siRNA）分子能够在细胞控制的许多方面起作用，它们能够反过来控制 DNA，停止基因的复制或者改变基因显现的水平。小 RNA 是一类 18～25 个碱基的双链 RNA，小 RNA 广义地定义为不直接参与蛋白质合成的 RNA。这些 RNA 分为三类：①加帽的小 RNA；②不加帽的小 RNA；③病毒小 RNA，具有调控基因表达的功能。小 RNA 分子通过互补序列能与 DNA 结合，抑制基因表达。20 世纪 90 年代初期研究发现小 RNA 能抑制植物的基因表达，随后在动物细胞中也发现了这一现象，双链 RNA 显著的抑制基因的表达，被称为 RNA 干扰（RNAi）。RNAi 证实了小 RNA 在基因沉默中发挥了作用，为遗传性疾病的治疗、病毒性疾病的治疗、研究基因功能、肿瘤治疗以及基因缺陷性动物制备提供了强大的工具。

RNAi 技术具有特异性和高效性，它是一种利用小的双链 RNA，高效、特异地阻断体内特定基因表达，促使 mRNA 降解，诱使细胞表现出特定基因缺失表型的技术。这种技术已经成为研究基因功能的重要工具，由于可以作为一种简单、有效的代替基因敲除的遗传工具，被“Science”评为 2001 年最重要的成果之一。目前，RNAi 的作用机制已经基本明确，并显示了广阔的应用前景。RNAi 技术作为新的基因阻断技术已经成功地应用于人类基因组功能的研究。

RNAi 具有以下几个重要的特征：①RNAi 是转录后水平的基因沉默机制；②RNAi 具有很高的特异性，只降解与之序列相应的单个内源基因的 mRNA；③RNAi 抑制基因表达具有很高的效率，表型可以达到缺失突变体表型的程度，而且相对很少量的 dsRNA 分子（数量远远少于内源 mRNA 的数量）就能完全抑制相应基因的表达，是以催化放大的方式进行的；④RNAi 抑制基因表达的效应可以穿过细胞界限，在不同细胞间长距离传递和维持信号，甚至传播至整个有机体以及可遗传等；⑤dsRNA 不得短于 21bp，并且长链 dsRNA 也在细胞内被 Dicer 酶切割为 21 bp 左右的 siRNA，并由 siRNA 来介导 mRNA 切割；大于 30bp 的 dsRNA 不能在哺乳动物中诱导特异的 RNA 干扰，而是细胞非特异性和全面的基因表达受抑和凋亡；⑥ATP 依赖性，在去除 ATP 的样品中 RNA 干扰现象降低或消失显示 RNA 干扰是一个 ATP 依赖的过程，可能是 Dicer 和 RISC 的酶切反应必须由 ATP 提供能量。

（一）RNAi 作用的机制

当病毒基因、人工转入基因、转座子等外源性基因随机整合到宿主细胞基因组内，并利用宿主细胞进行转录时，常产生一些 dsRNA。宿主细胞对这些 dsRNA 迅即产生反应，其胞质中的核酸内切酶 Dicer 将 dsRNA 切割成多个具有特定长度和结构的小片段 RNA（大约 21～23 bp），即 siRNA。siRNA 在细胞内 RNA 解旋酶的作用下解链成正义链和反义链，继之由反义 siRNA 再与体内一些酶（包括内切酶、外切酶、解旋酶等）结合形成 RNA 诱导的沉默复合物（RNA-induced silencing complex，RISC）。RISC 与外源性基因表达的 mRNA 的同源区进行特异性结合，RISC 具有核酸酶的功能，在结合部位切割 mRNA，切割位点即是与 siRNA 中反义链互补结合的两端。被切割后的断裂 mRNA 随即降解，从而诱发宿主细胞针对这些 mRNA 的降解反应。siRNA 不仅能引导 RISC 切割同源单链 mRNA，而且

可作为引物与靶 RNA 结合并在 RNA 聚合酶（RNA-dependentRNApolymerase，RdRP）作用下合成更多新的 dsRNA，新合成的 dsRNA 再由 Dicer 切割产生大量的次级 siRNA，从而使 RNAi 的作用进一步放大，最终将靶 mRNA 完全降解（图 15-1）。

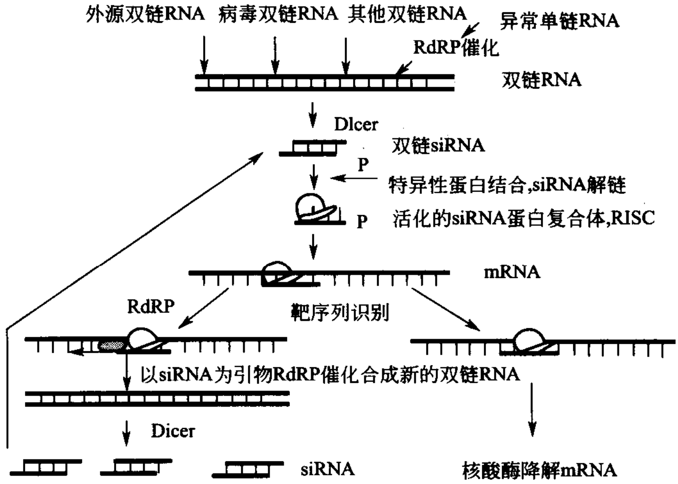


图 15-1　RNAi 作用机制示意图

siRNA 是 RNAi 的中心环节；它由两条包括 21 个核苷酸（21nt）的单链组成，其中有 19nt 形成双链；两条链 3′端各有 2nt 的对称性突出。这种对称性突出对于 siRNA 的功能是非常必要的。研究表明平端的 siRNA 不具有 RNAi 的功能。Bernstein 等率先成功地对参与第一步反应的核酸内切酶进行了提取与分析，将之命名为 Dicer。研究发现，Dicer 是 RNase Ⅲ家族中的一员，它含有一个螺旋酶结构域，一个 PAZ（pin wheel-argonaut-zwille）结构域；两个 RNase Ⅲ催化域及 dsRNA 结合结构域（dsRBD）。Dicer 通过 PAZ 结构域与其他相关蛋白的 PAZ 结构域结合，形成多亚单元复合物，特异性地将 dsRNA 切割为 21～23nt 的 siRNA。需要注意的是，编码 Dicer 的 Dicer 基因属于进化保守基因。参与第三步反应的酶，目前认为是一种与 Dicer 不同的核酸酶（RNase），至于这种核酸酶如何将靶 mRNA 特异性地降解，机理仍有待于进一步研究。Dicer 除了参与 RNAi 外，还能将 70nt 左右茎环状结构的 RNA 切割为 22nt 左右的片段。由于这些片段只能在特定的阶段和时间内表达，故称之为小时序 RNA（small temporal RNA，stRNA），stRNA 主要包括 lin-4 与 let-7。它们不编码蛋白质，却是调控线虫发育和成熟的重要调节因子。

（二）RNAi 实验方案

1. siRNA 的设计

人们发现小于 30bp 的 dsRNA 就可以有效引起基因沉默，并且不会产生非特异抑制现象，进一步研究表明抑制作用最强的是长 21bp、3′端有两个碱基突出的 siRNA。但 RNAi 技术要求 siRNA 反义链与靶基因序列之间严格的碱基配对，单个碱基错配就会大大降低沉默效应，而且 siRNA 还可以造成与其具同源性的其他基因沉默（也叫交叉沉默），所以在 siRNA 的设计中序列问题是至关重要的，参照靶向 mRNA 或 cDNA 序列设计 siRNA 是 RNA 干扰实验的关键。要求所设计的 siRNA 只能与靶基因具高度同源性而尽可能少地与其他基因同源。

siRNA 的设计应遵循以下原则：①以成熟的 mRNA 为标准，不应包含内含子序列。若以 DNA 序列为参照，则最好选取 cDNA 转录区＋50 至＋100 以后的下游某一段序列，两端

的非翻译区不作为 siRNA 设计的依据；②siRNA 二聚体的正义链和反义链以各含 21nt 为佳，3′为突出末端，即只有 19～20 个核苷酸组成碱基对；③ 3′突端为尿苷（U），5′凹端为腺苷（A）的 mRNA 序列应作为首选；具有 5′-AA-N_{19}-UU 模式，且鸟苷/胞苷（G/C）含量在 50% 左右的 mRNA 序列产生的效应最好；5′-AA-N_{21}或 5′-NA-N_{21}模式也具有较好的效应；在哺乳类动物细胞中进行 siRNA 表达载体的转染，其载体内 RNA 的转录要求 5′端第一个核苷为 G，因此 siRNA 的 5′端最好以 G 开始；④每个基因选择 4～5 个 siRNA 序列，然后进行 GenBank BLAST 查询，运用生物信息学方法进行同源性比较，剔除与其他基因有同源性的序列，选出一个特异性最强的 siRNA，确保所选 siRNA 编码不与其他基因同源。

2. siRNA 设计的基本原则

① siRNA 的长度一般在 21bp；

② 避免在使用距启动子或终止密码 50～100bp 内的互补序列；

③ 避免使用内含子互补序列；

④ 避免使用 4 个及以上相同碱基的重复序列，如 AAAA、CCCC；

⑤ 避免 GC 含量大于 60% 和小于 30% 的区域；

⑥ 避免重复和低复合体形成的序列；

⑦ 避免单核苷多态现象（SNP，Single Nucleotide Polymorphism）位点；

⑧ 进行 BLAST 检索，避免与其他的基因和序列出现结合；

⑨ 一定要有阴性对照，对照序列长度要与 siRNA 一致，且其碱基同 siRNA 相比最少要有 4～5 个碱基错配，并要保证对照序列不要与其他的基因有同源性。

3. Tuschl's 法则

① 在 cDNA 启动子下游 50～100bp 处开始选择目标序列；

② 首先寻找 23bp 末端为 AA（N_{19}），如果没有合适的序列，寻找 23bp 末端为 NA（N_{21}），并将 siRNA 的 3′端改为 TT，或寻找 NAR（N_{17}）YNN 的序列；

③ GC 含量在 50%左右。

A=Adenine；T=Thymine；R=Adenine 或 Guanine（Purines）；Y=Thymine 或 Cytosine（Pyrimidines）；N=Any

4. siRNA 设计的评估

Reynolds 评分法，评分越高，siRNA 的设计越合理。

标准	内　　容	评分	
		是	否
1	GC 的含量中到低(30%～52%)	1	
2	第 15～19 位核苷酸最少 3 个 A 或 U(有义链)	(每个 A 或 U)	
3	无内部重复(T_m①<20℃)	1	
4	19 位为 A(有义链)	1	
5	3 位为 A(有义链)	1	
6	10 位为 U(有义链)	1	
7	19 位不是 G 或 C(有义链)		−1
8	13 为不是 G(有义链)		−1

① $T_m=79.8+18.5\times\lg([Na^+])+(58.4\times GC\%/100)+[11.8\times(GC\%/100)^2]-(820/\text{序列的长度})$

例如，siRNA UUCUCCAGCUUCUAAAAUA 的 T_m

$T_m=79.8+18.5\times\lg(0.05)+(58.4\times31.6/100)+[11.8\times(31.6/100)^2]-(820/19)=32.19$

注：siRNA 的标准分为 6，高于 6 分均是合适的 siRNA。

（http：//boz094.ust.hk/RNAi/siRNA，该网站提供一种以 Excel 为模板的此原理 siRNA 的设计方法，另一种由 Dharmacon 公司提供）

5. siRNA 的合成

目前获得 siRNA 主要有三种方法，化学合成、体外酶法合成和体内转录。

(1) 化学合成　siRNA 可以直接进行化学合成，方便且不受碱基限制，但合成一个 siRNA 需先分别合成两条单链（每条链合成需 20 步反应，每步反应产物均需纯化），然后再退火成双链 RNA，合成耗时长且费用很高。目前有多家公司提供此项服务，如 Dhar macon 公司、Xeragon 公司、Proligo 公司和 Ambion 公司均可合成高质量的 siRNA。

(2) 体外酶法合成　与化学合成相比体外酶法合成要经济得多，此法需首先合成两段 29nt 的 DNA，包括 8 个与 T7 启动子 3′端互补的碱基（称为引导序列，leader sequence）和由 21 个碱基构成的 siRNA 编码序列。将这两段 DNA 分别和 T7 启动子混合，T7 启动子和 DNA 引导序列退火结合，然后用 DNA 聚合酶 Klenow 大片段补齐成为可用于转录的双链 DNA 模板，分别用 T7RNA 聚合酶进行体外转录，再将产物混合形成 dsRNA。新的双链 RNA 包含 5′端单链的引导序列和中间互补的 19 个碱基序列以及 3′端两个重复的 U (UU)，以 DNA 酶降解模板，同时用单链专一的核糖核酸酶（RNase）消化 5′的引导序列，由于 RNase 不能切开 U 碱基也不能降解双链 RNA，所以得到的产物就是所需的 21bp 的双链 siRNA——有 19 对碱基互补，3′端各有 2 个 U 突出。目前 Ambion 公司可提供相关试剂盒，用试剂盒可一次合成多个 siRNA 分子。用酶法合成的 siRNA 比化学合成的同序列 siRNA 活性高 20 倍，估计可能是由于 siRNA 纯度的差异造成的结果。

(3) 体内转录　目前已经构建了一些可以在真核细胞内表达 siRNA 的载体，包括被广泛应用的质粒载体和新研发的转染效率较高的病毒载体。质粒载体（以 Ambion 公司提供的 pSilencer 系列质粒为例，见图 15-2）都包含有一个 RNA 聚合酶Ⅲ（Pol Ⅲ）启动子和一个 4～5 个连续的 T 构成的转录终止位点以及选择标记，选用 Pol Ⅲ启动子是因为此启动子总是在距其一定距离的位置起始转录，而遇到 4～5 个连续的 T 就终止转录，并且终止于转录终止位点第 2 个碱基处，十分精确。此法需首先化学合成一段编码 siRNA 正义链和反义链的模板，正义序列与反义序列之间由一段环（loop）序列隔开，插入载体 Pol Ⅲ启动子下游，然后转入细胞，环序列可使转录出的 RNA 链折叠成具发夹结构的小 RNA 分子，这些小 RNA 可被细胞内的 Dicer 切割为长 21bp 的 siRNA——有 19 对互补碱基，3′端各有 2 个 U 突出，可引起特定基因沉默。与化学合成及体外酶法合成相比，表达载体是在细胞内通过转录持续产生 siRNA，所以可延长 siRNA 作用的时间，目前此法已被广泛应用，具体的技术细节可查看相关产品网站。

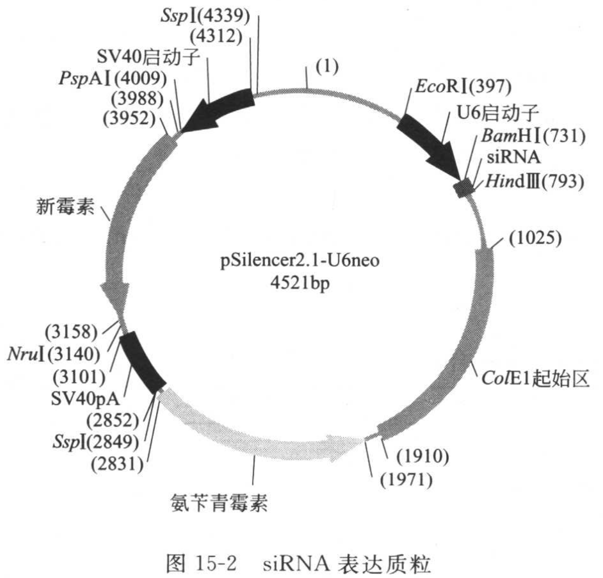


图 15-2　siRNA 表达质粒

(三) RNAi 敏感性的预测

不同实验室保留的同一细胞系或体外传代次数不同的细胞产生 RNAi 的效应不同，经过多次传代培养的细胞可能会丧失 RNA 干扰的能力；另外，来源于某种组织的特定细胞也可

能不支持现有 RNA 干扰的条件，前体细胞或干细胞接受 RNAi 不如分化成熟的细胞敏感。为确定细胞是否具有 RNAi 的效应，Elbash IR 等巧妙地构建了一种报告载体，为筛选细胞的 RNAi 效应提供了检测方案。将构建有荧光素酶基因质粒及阻断荧光素酶 mRNA 的 siRNA 共转染细胞，如果 siRNA 阻断了荧光素酶的表达，细胞裂解产物就不能使荧光底物发光，说明该细胞具有 RNAi 效应。

（四）siRNA 导入宿主细胞

在体外（*in vitro*）可用电穿孔法、脂质体介导法、病毒载体方法等将 siRNA 导入靶细胞，具体方法的选择应根据靶细胞的特性、试剂反应条件及实验室条件而定。一般来说，化学合成和体外酶法合成的 siRNA 可用电转移（electroporation）、微注射和转染的方法引入细胞，而表达质粒则常通过转染的方法导入靶细胞然后再表达 siRNA。为利于分析 RNA 干扰效应，导入的 siRNA 除包括目的基因外还应设立对照，如反向序列的 siRNA、中心区含有 1～ 2 个错配的 siRNA 或空白对照。

siRNA 的真核细胞内转染过程如下。

将合成 siRNA 转染到真核细胞为瞬间转染，如果要得到长期稳定干扰效应，可以将目标 cDNA 反向构建到条件表达载体，转染细胞后，转录生成 siRNA，达到干扰目的，并通过可以控制的条件，控制载体的转录。以下简要介绍使用脂质体将合成的 siRNA 转染至真核细胞内。

① 合成 siRNA 及对照 siRNA，两者的设计按上述原则进行。

② 制备 siRNA 复合物，70μl 2×退火缓冲液（200mmol/L 醋酸钾，4mmol/L 醋酸镁，60mmol/L HEPES-KOH，pH 7.4）中加入合成的 siRNA 和对照 siRNA，终浓度为 20μmol/L，加水至 140μl。

③ 90℃孵育 1min，然后 37℃孵育 1h。可将该混合液－20℃保存，以备后用。该混合液可以反复冻融，但要注意一定要放置于冰上，以减少 RNA 的水解。

④ 评估退火反应后复合物的完整性。用少许 0.5× TBE 缓冲液或蔗糖加样缓冲液稀释 1μl 20μmol/L siRNA 和对照 siRNA，0.5μl siRNA 复合物，进行 4% 琼脂糖凝胶电泳，0.5× TBE，80V，1h。紫外光下观察电泳条带，siRNA 复合物移动较慢，如条带出现拖尾，说明复合物形成不佳，重新进行反应。

⑤ 细胞培养的准备，用于转染的细胞要处在对数增殖期。

⑥ 进行转染前 24h，用 10ml 胰蛋白酶-EDTA 消化生长于 175ml 培养瓶中达 90% 融合的细胞，用无抗生素的培养基 1∶10 稀释，每孔（24 孔细胞培养板）加入细胞混悬液 0.5ml；如进行免疫荧光分析，细胞要生长在玻片上。

⑦ 5% CO_2，37℃培养至细胞达 50%融合。

⑧ 3μl 20μmol/L siRNA 复合物（相当 0.84μg，60pmol）同 50μl Opti-MEM1（life technologies）混合。

⑨ 在另一离心管中，将 3μl Oligofectamine（Invitrogen）或 4μl transIT-TKO（Mirus）同 12μl Opti-MEM1 混合，置室温 7～10min。

⑩ 将⑧的混合物加入⑨的混合物中，颠倒混匀，室温放置 20～25min，用于脂质复合物的形成，溶液变浑浊。

⑪ 加入 32μl Opti-MEM1 至 100μl，颠倒混匀。

⑫ 将 100μl 脂质复合物加入 50%融合的细胞中，并轻轻晃动培养板 30s；5% CO_2，

37℃培养 2～3d。

⑬ 2～3d 后，可通过 Western 印迹或免疫荧光分析进行 siRNA 干扰的检测。

（五）结果分析

（1）形态学观察　siRNA 导入宿主细胞后，其生物学效应一般在 1～3d 内出现，但依转染的效率、目的基因表达丰度或其蛋白产物半寿期的长短而略有差异。若目的基因与细胞生长分化有关，则细胞形态学改变可通过相差显微镜、荧光显微镜，甚至光学显微镜观察到。

（2）免疫荧光　原位状态下分析目的基因表达的蛋白质分布情况和量的变化。被测蛋白质进行免疫荧光标记后，可以在荧光显微镜或激光扫描共聚焦显微镜下进行观察与定性分析。

（3）Western 印迹分析　主要步骤如下：细胞分离、提取、漂洗；煮沸法碎裂细胞并提取蛋白组分；SDS-PAGE 电泳分离蛋白质；将蛋白质转移至硝酸纤维素膜上；相应抗体与蛋白质在硝酸纤维素膜上进行孵育后成像分析。

（4）RT-PCR 检测目的 mRNA　当无法得到目的基因的蛋白产物的特异性抗体时，可应用 RT-PCR 技术对基因阻断前后的 mRNA 表达进行分析。

（六）RNAi 在基因功能分析中的应用

RNAi 技术已经作为一种新型的反向遗传学技术，成为基因功能研究的重要工具。在基因的研究中，首先利用此技术剔除待研究的目标基因，然后通过细胞表型改变的观察、免疫蛋白印迹、免疫荧光等方法即可对基因的表达及功能进行分析。

作为一种新的、快速的解读基因的功能和基因的表达机制以及基因之间的相互关系的方法，RNAi 技术在基因功能研究上有其独特的优点：①简单易行，容易开展；②与基因敲除（gene knockout）相比实验周期短，成本低；③与反义技术相比具有高度特异性和高效性；④可进行高通量（high-throughout）基因功能分析。RNAi 技术的诸多优点使它很快便被作为功能基因组学研究的主要方法。RNAi 将成为研究基因功能不可或缺的工具。

总之，RNAi 是一个以小分子 RNA 为中心的真核细胞基因表达调控系统，它可以在多个层面调节基因表达和细胞的增殖分化，而且 RNAi 不仅发生在 DNA 转录后的 mRNA 水平上，siRNA 亦可进入细胞核内指导相应 DNA 的合成后修饰，从而在 DNA 转录水平上调节基因的表达。RNAi 技术作为一种新的基因阻断技术方兴未艾，已经成为人类基因功能研究的得力工具，使人们能更迅速、更准确地剖析基因的功能，分析基因之间错综复杂的联系和相互作用。

第三节　核酶技术

20 世纪 80 年代初期，美国科学家 Cech 和 Altman 分别发现了具有生物催化功能的 RNA，即核酶（ribozyme）。核酶是一类具有酶特性的 RNA 分子，通过催化靶位点 RNA 链中磷酸二酯键的断裂破坏 mRNA 的状态，从而阻断该基因的表达。核酶由于其具有较为

稳定的空间结构、可重复使用、具有封闭和切割双重作用等特点而被誉为“基因剪刀”，越来越广泛地应用于基因治疗等领域。

1. 核酶的结构

核酶的基本组成中应有中间极端保守的核苷酸序列（活性中心）和两端的引导序列，当两端引导序列与靶 RNA 互补结合时，中间极端保守序列即在该位点切断。

常见的核酶有锤头状核酶（hammerhead ribozyme）、发夹状核酶（hairpin ribozyme）和斧头状核酶（hatchety ribozyme）三种。因为锤头状核酶分子较小，结构简单，RNA 酶各位点的序列限制少，在 RNA 中的出现率较多，且核酶催化中心的保守序列业已确定，所以人工设计容易，应用较多。

2. 核酶的设计与合成

核酶的设计目的在于选择特异性高、活性强、性质稳定的核酶。一般来讲，要获得特异性高的核酶就必须找到有效的靶序列，要取得较好活性和稳定性就需要对核酶进行修饰。所以，下面将从两方面介绍核酸的设计与合成。

（1）有效靶序列的选择　最适靶位受到相当多因素影响，经验方法是唯一的途径来鉴别最有效位点，而这需要检测每一可能的靶位。这里介绍一种核酶库技术，其基本原理是先合成与催化区中心部分序列重叠的单链寡核苷酸引物，然后将其与三磷酸脱氧核苷和 *Taq* 聚合酶混合孵育，所有的核苷酸以相同的概率随机结合于其侧翼，这样就产生了一系列带有随机侧链序列的核酶。将其用特异性限制性内切酶剪切，并将其克隆到载体相应的位点后，就可生产大量不同侧链区的核酶，即核酶库。若侧链区的最大长度为 26nt（两侧各为 13），则理论上可获得 10^9 个不同序列的质粒克隆。最后用特异性位点对这些核酶进行鉴定。由于这种核酶库提供了足够多序列不同的核酶，大大提高了找到有效核酶的可能。

（2）核酶的修饰　未经修饰的核酶在细胞内迅速被核酸酶降解（半衰期大约只有 0.1min），所以需要对核酶进行修饰。但至今尚无完整的理论可以用于指导核酶的修饰，大量的试验仍是依赖于经验。对核酶的修饰可以归为三类，即磷酸二酯骨架的修饰、戊糖环的修饰和碱基的修饰。

3. 核酶的导入

有两种途径可以把核酶导入到细胞内以取得抑制基因成功的表达。一种是外源性导入，即把预先设计合成的核酶通过细胞内注射法、电穿孔法、氯化钙或脂质体介导法技术直接导入细胞内发挥作用。另一种是内源性导入，这种方法是将编码核酶的基因（DNA）构建于表达载体上再导入细胞，然后再转录产生大量核酶分子发挥切割作用。后一种基因转移方法又可分为两种：①病毒介导法利用逆转录病毒、腺病毒相关病毒作为转移载体，是近年来应用较广的导入方法，缺点是整合位点不稳定，转染率不高；②物理-化学法，是把质粒 DNA 复合体或包裹在微小金粒外的 DNA 通过受体介导法、气雾剂法、颗粒轰击法直接注入体内的一类方法。

4. 核酶的应用

（1）在癌基因研究中的应用　利用核酶独特的切割失控基因或有害基因的分子剪刀方式，有可能阻断癌基因的表达，从而逆转肿瘤细胞恶性增殖，并诱导其分化和凋亡。1992 年 Kashami 等采用 *ras* 癌基因特异的核酶逆转入膀胱癌细胞系 EJ 细胞的恶性表型，肿瘤细胞的致病性和转移性均降低。

（2）在端粒酶研究中的应用　在高等生物细胞中，除了精原细胞和少数造血细胞存在端

粒酶活性外，其余体细胞的端粒酶均处于失活状态。而当细胞恶变时，端粒酶则被激活，使端粒酶长度不随细胞分裂而丢失，从而使细胞逃避老化死亡，保持无限复制和增殖。1995年端粒酶的组分被克隆以后，人们开始尝试用反义技术作用于端粒酶 RNA 从而抑制端粒酶活性。Feng 等构建了表达反义 RNA 的真核表达质粒，导入细胞使细胞端粒酶活性降低至原来的水平，细胞传至 26 代后出现凋亡。

（3）在多药耐药性（multidrug resistance，MDR）研究中的应用　MDR 是指肿瘤细胞对结构和功能并不相同的化疗药物产生的交叉耐受，是肿瘤化疗失败的重要原因。细胞水平的研究资料表明，单独或联合应用钙通道阻滞剂、抗生素、孕激素、环孢素和单克隆抗体等药物，能够提高耐药细胞的敏感性，逆转 MDR。然而，临床上应用这些药物逆转 MDR，其主要困难是机体难以耐受有效治疗浓度。核酶的发现为逆转 MDR 的基础研究和临床应用开辟了新的领域。

（4）在病毒感染的基因治疗中的应用　既然核酶可专一性地切割 RNA 靶分子，应用核酶技术阻断病毒复制和基因表达就成为可能。由于核酶是在 RNA 水平上发挥作用，通过特异性与靶 RNA 配对并切割从而阻断了病毒的繁殖与复制，并且长期使用对细胞无副作用，其结构简单，可以人工设计，因此很有希望用于病毒感染的基因治疗。国内各研究机构已经分别开展了针对乙型、丙型、丁型肝炎病毒核酶的研究。

5. 问题与展望

核酶对基因表达的抑制具有特异性强、作用敏感而迅速等优点，目前在体内外的实验都取得了较为满意的结果。因此，它极有可能发展为基因治疗的药物或者研究基因功能的工具。但就目前的研究水平，距临床使用还有一定的距离，主要的问题在于核酶的转导系统、核酶基因表达、切割部位和切割效率等几个方面。

第四节　RNA 错折叠（ODMiR）技术

常规的反义 RNA 技术是设计出与靶 RNA 互补的寡核苷酸，通过它与靶 RNA 结合来干扰转录或翻译过程，使蛋白质不能表达。2002 年，Childs 等发明了一种新型的反义 RNA 技术——寡核苷酸导向的 RNA 错折叠技术，它将反义 RNA 技术进一步简化，从另一个角度丰富了反义 RNA 技术。

1. RNA 错折叠的原理

常规的反义 RNA 技术是通过把互补的 RNA 分子附加到靶 RNA 链上以阻断靶 RNA 的表达，而 RNA 错折叠技术则比较巧妙地通过附加一小段寡核苷酸片段来改变折叠 RNA 的动力学，导致 RNA 的结构和形状发生改变，最终封闭靶 RNA 的翻译，这种方法被称为寡核苷酸导向的 RNA 错折叠（oligonucleotide directed misfolding of RNA，ODMiR）。

RNA 携带的是基因中 DNA 所编码的遗传指令，所以它成为基因治疗中的一个重要靶分子。RNA 结构虽然迂回曲折，环结交错，但是一般来说，RNA 能够行使其功能要具备以下两个条件：①折叠成二级结构或三级结构；②折叠成的结构要具有最低的自由能。如果这个结构被破坏，RNA 则不能正常地进行转录和翻译过程。根据该理论基础，罗切斯特大

学医学研究中心的科学家们设计了一小段寡核苷酸，把它附加到靶RNA上，使靶RNA的结构改变，导致其失去活性。该实验以白色念球菌为材料，通过设计或筛选出的寡核苷酸与靶RNA结合后造成靶RNA的错误折叠，抑制了Ⅰ型内含子的剪接。白色念球菌的Ⅰ型内含子可能呈现出不同的形状，常见的一种是最适结构，另一种是次适结构，它们的自由能只相差2.2kcal/mol，结构差异在于次适结构在252～259位核苷酸处出现一个5×8的环，这个环对全长Ⅰ型内含子的折叠有限制作用，只要设计出一段与252～259核苷酸序列互补的反义RNA片段来封闭这段核苷酸，就可以使这一环状结构保持稳定，从而改变Ⅰ型内含子的动力学，抑制该内含子的剪接。还应考虑的一个影响因素是所使用的反义寡核苷酸的浓度，不同浓度下的反义寡核苷酸对Ⅰ型内含子剪接的抑制程度不同。L（TACCTTTC）是与252～259位核苷酸互补的一小段寡核苷酸，在其浓度为150nmol/L时可以抑制50%的内含子进行自我剪接，而在浓度为100μmol/L时则没有控制作用。另外，在实验中还发现有些并不与靶RNA序列互补，但是碱基组成与L（TACCTTTC）相同的小寡核苷酸对内含子的剪接也具抑制作用，如L（CCTTATCT）。由此可见小寡核苷酸的碱基排列顺序对靶RNA的作用影响不大。

2. RNA错折叠的前景

寡核苷酸导向的RNA错折叠技术虽然是以白色念球菌为试验对象发明的，但却具有普遍性，同时对小寡核苷酸与靶RNA之间的分子识别原理已经了解比较深入，小寡核苷酸片段也非常容易获得，小寡核苷酸中碱基序列的顺序与靶RNA的动力学特征关系不大，这些特点都使新药的开发过程变得更加简单和经济。该技术中所使用的反义RNA比常规反义技术中所使用的RNA要短得多，只需8～12个核苷酸就能破坏一个长达400个核苷酸的RNA分子。这对于代价昂贵的反义技术来说，不仅节省了一大笔费用，引起的副作用也比较小。反义RNA技术成功地抑制了疱疹病毒、流感病毒、HIV、肉瘤病毒等对组织细胞的侵袭，而寡核苷酸导向的RNA错折叠技术开辟了反义RNA技术发展的新方向。虽然RNA错折叠技术目前还不是很成熟，仍处于基础研究阶段，在不久的将来，RNA错折叠技术有望成为治疗肿瘤，病毒、细菌传染性疾病等的重要手段之一。

（杨仕林　伍欣星）

附　　录

一、常用试剂

1. 常用缓冲液

TE	pH7.4	10mmol/L Tris・HCl(pH7.4),1mmol/L EDTA(pH8.0)	
	pH7.6	10mmol/L Tris・HCl(pH7.6),1mmol/L EDTA(pH8.0)	
	pH8.0	10mmol/L Tris・HCl(pH8.0),1mmol/L EDTA(pH8.0)	
Tris buffer	2mol/L pH7.4		
	水	850ml	
	Tris base	242.2g	
	小心缓慢加入 HCl,不断搅动,至 pH7.4		
TES	配 100ml		
	Tris 20mmol/L pH7.4	1.0ml 2mol/L 贮存液	
	EDTA・Na_2 0.1mmol/L	20.0μl 0.5mol/L 贮存液	
	NaCl 10mmol/L	0.2 ml 5mol/L 贮存液	
	加水至	100.0ml	
TMG	配 100ml		
	10mmol/L Tris・HCl(pH7.4)	0.5ml 2mol/L 贮存液	
	10mmol/L $MgSO_4$	1.0ml 1mol/L 贮存液	
	0.01%白明胶	10.0mg	
	加水至	100.0ml	
STE	(TEN)		
	0.1mol/L NaCl		
	10mmol/L Tris・HCl(pH8.4)		
	1mmol/L EDTA(pH8.0)		
STET	0.1mol/L NaCl		
	10mmol/L Tris・HCl(pH8.0)		
	1mmol/L EDTA(pH8.0)		
	5% Triton X-100		
TNT	10mmol/L Tris・HCl(pH8.0)		
	150mmol/L NaCl		
	0.05%吐温-20		
SSC	3mol/L NaCl	配 1L	配 4L
	柠檬酸三钠・H_2O 0.3mol/L	175.3g	701.2g
	水	88.2g	352.9g
	1mol/L HCl 调 pH 值至 7.0	800.0ml	3200ml
		加水到 1L	加水到 4L
SSPE	20×配 1L		
	3mol/L NaCl	175.3g	
	Na_2HPO_4・H_2O 0.2mol/L	27.6g	
	EDTA-Na_2 0.02mol/L	加 40ml 0.5mol/L EDTA 贮存液	
	加水	800ml	
	用 10mol/L NaOH 调 pH 值至 7.4 加水至 1L		

2. 电泳缓冲液

(1) 测序凝胶加样缓冲液

98%	去离子甲酰胺
10mmol/L	EDTA
0.025%	二甲苯青 FF
0.025%	溴酚蓝

甲酰胺：许多批号的试剂级甲酰胺，其纯度符合要求，无需再进行处理，不过一旦呈黄色，应在磁力搅拌器上将甲酰胺与 Dowex XG8 混合床树脂共同搅拌 1h 进行去离子处理，再用 Whatman 1 号滤纸滤过两次，去离子甲酰胺分装成小份，充氮贮存于－70℃。

(2) 常用的电泳缓冲液

缓冲液	使用液	浓贮存液/L
Tris-乙酸(TAE)	1×： 0.04mol/L Tris-乙酸 0.001mol/L EDTA	50×： 242g Tris 碱 57.1ml 冰醋酸 100ml 0.5mol/L EDTA(pH8.0)
Tris-磷酸(TPE)	1×： 0.09mol/L Tris-磷酸 0.002mol/L EDTA	10×： 108g Tris 碱 15.5ml 85%磷酸(1.679g/ml) 40ml 0.5mol/L EDTA(pH8.0)
Tris-硼酸(TBE)①	0.5×： 0.045mol/L Tris-硼酸 0.001mol/L EDTA	5×： 54g Tris 碱 27.5 硼酸 20ml 0.5mol/L EDTA(pH8.0)
碱性缓冲液②	1×： 50mmol/L NaOH 1mmol/L EDTA	1×： 5ml 10mol/L NaOH 2ml 0.5mmol/L EDTA(pH8.0)
Tris-甘氨酸③	1×： 25mmol/L Tris 250mmol/L 甘氨酸 0.1% SDS	5×： 15.1g Tris 碱 94g 甘氨酸(电泳级)(pH8.3) 50ml 10% SDS(电泳级)

① 为避免 TBE 浓溶液长时间存放后会形成沉淀物，可在室温下用玻璃瓶保存 5×溶液，若出现沉淀后则应废弃。以往都以 1×TBE 作为使用液（即 1∶5 稀释浓贮液）进行琼脂糖凝胶电泳，但 0.5×的使用液已具备足够的缓冲容量。目前几乎所有的琼脂糖胶电泳都以 1∶10 稀释的贮存液作为使用液。进行聚丙烯酰胺凝胶电泳使用的是 1×TBE，是琼脂糖凝胶电泳时使用液浓度的 2 倍。这是由于聚丙烯酰胺凝胶垂直槽的缓冲液槽较小，需加强离子强度，提供稍高一些的电流。

② 碱性电泳缓冲液应现用现配。

③ Tris-甘氨酸缓冲液用于 SDS-聚丙烯酰胺凝胶电泳。

3. 2×SDS 凝胶加样缓冲液

100mmol/L	Tris·HCl
200mmol/L	二硫苏糖醇(DTT)
4%	SDS(电泳级)
0.2%	溴酚蓝
20%	甘油

不含二硫苏糖醇（DTT）的 2×SDS 凝胶加样液可保存于室温，应在临用前取 1mol/L 二硫苏糖醇贮存液现加于上述缓冲液中。

4. 凝胶加样缓冲液

缓冲液类型	6×缓冲液	贮存温度
Ⅰ	0.25%溴酚蓝 0.25%二甲苯青 FF 40g/100ml 蔗糖水溶液	4℃
Ⅱ	0.25 溴酚蓝 0.25%二甲苯青 FF 15%聚蔗糖(Ficoll 400 型)水溶液	室温
Ⅲ	0.25%溴酚蓝 0.25%二甲苯青 FF 30%甘油水溶液	4℃
Ⅳ	0.25%溴酚蓝 40g/100ml 蔗糖水溶液	4℃
Ⅴ	碱性加样缓冲液： 300mmol/L NaOH 6mmol/L EDTA 18%聚蔗糖(Ficoll 400 型)水溶液 0.15%溴甲酚绿 0.25%二甲苯青 FF	4℃

使用以上凝胶加样缓冲液的目的有：①增大样品密度，以确保 DNA 均匀进入样品孔内。使样品呈现颜色，从而使加样操作更为便利。含有在电场中能以可预知速率向阳极泳动的染料。溴酚蓝在琼脂糖中移动的速率约为二甲苯青的 2.2 倍，而与琼脂糖浓度无关。②以 0.5×TBE 作电泳液时，溴酚蓝在琼脂糖中的泳动速率约与长 300bp 的双链线状 DNA 相同，而二甲苯青 FF 的泳动则与长 4kb 的双链线状 DNA 相同。在琼脂糖浓度为 0.5%～1.4%的范围内，这些对应关系受凝胶浓度变化的影响并不显著。③对于碱性凝胶应当使用溴甲酚绿作为示踪染料，因为在碱性 pH 条件下其显色较溴酚蓝更鲜明。

5. 杂交缓冲液

① 杂交缓冲液-N（用于切口平移探针），配约 800ml

dextran sulfite（硫酸葡聚糖）	80g
H_2O	300ml
甲酰胺（去离子）	320ml
20×SSC 贮存液	160.0ml
Tris 2mol/L（pH7.4）	8.0ml
100× Denhardt's 液	8.0ml
鲑鱼精子贮存液（10mg/ml）	2.2ml

可搅动过夜，Whatman 滤纸滤过，－20℃可存数月

② 杂交缓冲液-S（用于合成探针）

10% PVP（聚乙烯吡咯烷酮）	3ml
10% BSA	3ml
10% Ficoll	3ml
鲑鱼精子贮存液（10mg/ml）	2.2ml
20×SSC 贮存液	15.0ml
酵母菌 RNA	250.0mg
加水至	150ml

③ 连接酶缓冲液，10×，用于 T4 DNA 连接酶，配 2ml

Tris 0.5mol/L pH7.4	0.5ml 2mol/L 贮存液
$MgCl_2$ 0.1mol/L	0.2ml 1mol/L 贮存液

DTT 0.2mol/L	0.4ml 1mol/L 贮存液
ATP 10mmol/L	0.2ml 100mmol/L 贮存液，用 NaOH 调至 pH7.0
BSA 50μg/ml	100μg
加水至	2ml

④ SSPE 缓冲液，20×，(用于噬菌体操作)，配 1L

NaCl (3mol/L) 175.3g

$NaH_2PO_4 \cdot H_2O$ (0.2mol/L) 27.6g

EDTA·Na_2 (0.2mol/L) 加 40ml 0.5mol/L EDTA

加水至 800ml

用 10mol/L NaOH 调至 pH7.4，高压蒸汽灭菌

6. 常用贮存液

溶液	配制方法
30%丙烯酰胺溶液	将 29g 丙烯酰胺和 1g *N*,*N'*-亚甲双丙烯酰胺溶于总体积为 60ml 的水中。加热至 37℃溶解之，加水至终体积为 100ml。用 0.45μm 滤膜过滤除菌，并保证该溶液的 pH 应不大于 7.0，置棕色瓶中保存于室温。注意：丙烯酰胺和 *N*,*N'*-亚甲双丙烯酰胺具有很强的神经毒性并可以通过皮肤吸收，其作用具有累积性。称量丙烯酰胺和亚甲双丙烯酰胺时应戴手套和面具。聚丙烯酰胺无毒，但也应谨慎操作，因为它还可能会含有少量未聚合丙烯酰胺。一些价格较低的丙烯酰胺和亚甲双丙烯酰胺通常含有一些金属离子，在丙烯酰胺贮存液中加入大约 0.2 倍体积的单床混合树脂(MB-1Mallinckrodt)，搅拌过夜，然后用 Whatman 1 号滤纸过滤以纯化之。在贮存期间，丙烯酰胺和亚甲双丙烯酰胺会因为光或碱的催化缓慢转化成丙烯酰和亚甲双丙烯酸。因此，应将该溶液装入棕色瓶中，并于 4℃保存
0.1mol/L 腺苷三磷酸(ATP)	在 0.8ml 水中溶解 60mg ATP，用 0.1mol/L NaOH 调至 pH7.0，用蒸馏水定容至 1ml，分装成小份保存于－70℃
10mol/L 乙酸胺	把 770g 乙酸胺溶解于 800ml 水中，加水定容至 1L 后过滤除菌
10g/100ml 过硫酸铵	把 1g 过硫酸铵溶解于终量为 10ml 的水溶液中，该溶液可在 4℃保存数周
BCIP	把 0.5g 5-溴-4-氯-3-吲哚磷酸二钠盐(BCIP)溶解于 10ml 100%的二甲基甲酰胺中，保存于 4℃
2×BES 缓冲盐溶液	用总体积 90ml 的蒸馏水溶解 1.07g BES[*N*,*N*-双(2-羟乙基)-2-氨基乙磺酸]、1.6g NaCl 和 0.027g Na_2HPO_4，室温下用 HCl 调节该溶液的 pH 值至 6.96，然后加入蒸馏水定容至 100ml，用 0.22μm 滤膜过滤除菌，分装成小份，保存于－20℃
1mol/L $CaCl_2$	在 200ml 纯水中(Milli-Q 水或相当级别的水)溶解 54g $CaCl_2 \cdot 6H_2O$，用 0.22μm 滤膜过滤除菌，分装成小份，保存于－20℃ 注意：制备感受态细胞时，取出 1 小份解冻并用纯水稀释至 100ml，用 Nalgene(0.45μm 孔径)过滤除菌，然后骤冷至 0℃
2.5mol/L $CaCl_2$	在 20ml 蒸馏水中溶解 13.5g $CaCl_2 \cdot 6H_2O$，用 0.22μm 滤器过滤除菌，分装成 10ml 小份，贮存于－20℃
1mol/L 二硫苏糖醇(DTT)	用 20ml 0.01mol/L 乙酸钠溶液(pH5.2)，溶解 3.09g DTT，过滤除菌后分装成 1ml 小份，贮存于－20℃。注意：DTT 或含有 DTT 的溶液不能进行高压处理
0.5mol/L EDTA (pH8.0)溶液	在 800ml 水中加入 186.1g 二水乙二胺四乙酸二钠(EDTA-Na·$2H_2O$)，在磁力搅拌器上剧烈搅拌，用 NaOH 调节溶液的 pH 值至 8.0(约需 20g NaOH 颗粒)，然后定容至 1L，分装后高压灭菌备用。注意：EDTA 二钠盐需加入 NaOH 将溶液的 pH 值调至接近 8.0 时，才能完全溶解
溴化乙锭(10mg/ml)	在 100ml 水中加入 1g 溴化乙锭，磁力搅拌数小时以确保其完全溶解，然后用铝箔包裹容器或转移至棕色瓶中，保存于室温。小心：溴化乙锭是强诱变剂并有中度毒性，使用含有这种染料的溶液时务必戴上手套，称量染料时要戴面罩。这些溶液使用后应进行净化处理
2×HEPES	用总量为 90ml 的蒸馏水溶解 1.6g NaCl、0.074g KCl、0.027g $Na_2PO_4 \cdot 2H_2O$、0.2g 葡聚糖和 1g HEPES，用 0.5mol/L NaOH 调节 pH 值至 7.05，再用蒸馏水定容至 100ml。用 0.22μm 滤器过滤除菌，分装成 5ml 小份，贮存于－20℃
IPTG	IPTG 为异丙基硫代-β-D-半乳糖苷(分子量为 236.4)，在 8ml 蒸馏水中溶解 2g IPTG 后，用蒸馏水定容至 10ml，用 0.22μm 滤器过滤除菌，分装成 1ml 小份，贮存于－20℃

续表

溶液	配制方法
1mol/L 乙酸镁	在800ml水中溶解214.46g四水乙酸镁,用水定容至1L,过滤除菌
1mol/L $MgCl_2$	在800ml水中溶解203.3g $MgCl_2 \cdot 6H_2O$,用水定容至1L,分装成小份并高压灭菌备用。注意:$MgCl_2$ 极易潮解,应选购小瓶(如100g)试剂,启用新瓶后勿长期存放
10%十二烷基硫酸钠(SDS)	在900ml水中溶解100g电泳级SDS,加热至68℃助溶,加入几滴浓盐酸,调节溶液的pH值至7.2,加水定容至1L,分装备用。注意:SDS的微细晶粒易于扩散,因此称量时要戴面罩,称量完毕后要清除残留在称量工作区和天平上的SDS,10%SDS溶液无需灭菌
20×SSC	在800ml水中溶解175.3g NaCl和88.2g柠檬酸钠,加入数滴10mol/L NaOH溶液调节pH值至7.0,加水定容至1L,分装后高压灭菌
20×SSPE	在800ml水中溶解175.3g NaCl、27.6g $NaH_2PO_4 \cdot H_2O$ 和7.4g EDTA,用NaOH溶液调节pH值至7.4(约需6.5ml 10ml/L NaOH),加水定容至1L,分装后高压灭菌
100%三氯乙酸溶液	在装有500g TCA的瓶中加入227ml水,形成的溶液含有100g/100ml TCA
1mol/L Tris	在800ml水中溶解121.1g Tris碱,加入浓HCl调节pH至所需值:pH7.4,HCl 70ml;pH7.6,HCl 60ml;pH8.0,HCl 42ml。应使溶液冷至室温后,方可最后调定pH,加水定容至1L,分装后高压灭菌。注意:如1mol/L溶液呈现黄色,应予丢弃,并置备质量更好的Tris。尽管多种类型的电极均不能准确测量Tris溶液的pH,但仍可向大多数厂商购得合适的电极。Tris溶液的pH因温度而异,温度每升高1℃,pH大约降低0.03个单位。例如,0.05mol/L的溶液在5℃、25℃和37℃时的pH分别为9.5、8.9、8.6
Tris缓冲盐溶液(TBS)(25mmol/l Tris)	在800ml蒸馏水中溶解8g NaCl、0.2g KCl和3g Tris碱,加入0.015g苯酚并用HCl调至pH值至7.4,用蒸馏水定容至1L,分装后在 1.034×10^5Pa高压下蒸汽灭菌20min,于室温保存
X-gal	X-gal为5-溴-4-氯-3-吲哚-β-D半乳糖苷。用二甲基甲酰胺溶解X-gal,配制成20mg/ml的贮存液。保存于一玻璃管或聚丙烯管中,装有X-gal溶液的试管须用铝箔封裹以防因受光照而被破坏,并应贮存于-20℃。X-gal溶液无需过滤除菌
磷酸盐缓冲溶液(PBS)	在800ml蒸馏水中溶解8g NaCl、0.2g KCl、1.44g Na_2HPO_4 和0.24g KH_2PO_4,用HCl调节溶液的pH值至7.4,加水定容至1L,在 1.034×10^5Pa高压下蒸汽灭菌20min。保存于室温
1mol/L 乙酸钾(pH7.5)	将9.82g乙酸钾溶解于90ml纯水(Milli-Q水或相当级别的水)中,用2mol/L乙酸调节pH值至7.5后,加入纯水定容到1L,保存于-20℃
乙酸钾溶液(用于碱裂解)	在60ml 5mol/L乙酸钾溶液中加入11.5ml冰醋酸和28.5ml水,即成钾浓度为3mol/L而乙酸根浓度为5mol/L的溶液
3mol/L 乙酸钠	在800ml水中溶解408.1g三水乙酸钠,用冰醋酸调节pH值至5.2,或用稀乙酸调节pH值至7.0,加水定容到1L,分装后高压灭菌
5mol/L NaCl	在800ml水中溶解292.2g NaCl,加水定容至1L,分装后高压灭菌

7. 杂交实验所用封闭剂

试剂	用途	配制方法
Denhardt试剂	Northern杂交 使用RNA探针的杂交;单拷贝序列的Southern杂交;将DNA固定于尼龙膜上的杂交	Denhardt试剂通常配制成50×贮存液,过滤后保存于-20℃。可将该贮存液10倍稀释于杂交液(常为含有0.5%SDS和100g/ml经变性并被打断的鲑鱼精DNA的6×SSC或6×SSPE)中。50×Denhardt液中含5g聚蔗糖(Ficoll,400型)、5g聚乙烯吡咯烷酮和5g牛血清白蛋白(组分V),加水至终体积为500ml
BLOTTO	Grunstein-Hogness杂交 Benton-Davis杂交;除单拷贝序列Southern杂交以外的所有Southern杂交斑点印迹	1×BLOTTO(牛乳转移技术优化液)是含5%脱脂奶粉和0.02%叠氮钠的水溶液,应保存于4℃。使用前可用杂交液稀释25倍。BLOTTO不应与高浓度的SDS并用,因为后者会导致牛奶中的蛋白质析出。如果杂交背景不合要求,可在杂交液中加入NP-40至终浓度为1%。BLOTTO不能用做Northern杂交的封闭剂,因为这一封闭剂中可能含有RNA酶。注意:叠氮钠有毒性,取用时需待手套小心操作。含叠氮钠的溶液应予以明确标记

续表

试剂	用途	配制方法
肝素	Southern 杂交 原位杂交	肝素(Sigma,从猪中提取的二级产品或相当等级的产品)用 4×SSPE 或 4×SSC 溶解配制成 50mg/ml 的浓度,保存于 4℃。肝素在含有葡聚糖硫酸酯的杂交液中用做封闭剂的浓度为 500μg/ml,在不含葡聚糖硫酸酯的杂交液中的浓度为 50μg/ml。
鲑鱼精 DNA 10mg/ml	Southern、Northern 杂交	称取鲑鱼精 DNA(S,Ⅲ型,钠盐)10mg 放入灭菌管中,加入 100mmol/L Tris·HCl,5mmol/L EDTA 缓冲液 1ml,用超声波处理,直至液体透亮,完全溶解,小份分装,贮存于−20℃。临用前置沸水浴中加热 5min,然后迅速在冰浴中骤冷。于杂交液中应含有 100μg/ml 经变性并打断的鲑鱼精 DNA

8. 铬酸洗液及配制方法

铬酸洗液是实验室中最常用的强氧化洗液之一，一般称之为铬酸混合剂（chromic acid mixture）或洗液（clean mixture，clean solution）。重铬酸钾为氧化剂，它与硫酸所组成的铬酸洗液氧化力强。铬酸洗液的配制方法虽然十分简单，但是反应过程中能产生大量的热，并有迸溅的危险，所以在配制铬酸洗液时要特别注意。新配制的铬酸洗液应为深红色，配方中浓硫酸含量高的洗涤效果好。

一些洗液的配方

重铬酸钾/g	水/ml	浓硫酸/ml	重铬酸钾/g	水/ml	浓硫酸/ml
10	100	30	15	—	500
10	100	75	50	—	1000
10	100	100	50	5	1000
12.7	100	900(18)	80	1000	100
12.7	100	100	200	500	500
2	5	65			

铬酸洗液的配制应视需要来选择，无论使用哪种配方，只要是含水的，即使含水量很少，也要先将重铬酸钾溶于水中，必要时可加热促溶，待溶液冷却后，在不断搅拌下将浓硫酸缓缓加入，注意不要出现铬酸结晶。配制铬酸洗液应使用耐热的容器，如酸缸等。在大量配制时要穿戴必要的防护用品。

二、常用参数与数据

1. 常用核酸的长度和分子量

核酸	长度	分子量	核酸	长度	分子量
λ DNA	48502bp(双链环状)	3.0×10^7	18S rRNA	1900base	6.1×10^5
pBR322	4363bp(双链)	2.8×10^6	19S rRNA	1700base	5.5×10^5
28S rRNA	4800base	1.6×10^6	5S rRNA	120base	3.6×10^4
23S rRNA	3700base	1.2×10^6	tRNA(大肠杆菌)	75base	2.5×10^4

2. 常用核酸、蛋白质换算数据

① 质量换算

$1\mu g=10^{-6}g$

$1ng=10^{-9}g$

$1pg=10^{-12}g$

$1fg=10^{-15}g$

② 分光光度换算

$1A_{260nm}$双链 DNA＝50μg/ml DNA

$1A_{260nm}$单链 DNA＝33μg/ml DNA

$1A_{260nm}$单链 RNA=40μg/ml DNA

③ DNA 摩尔换算

1μg 1000bp 长度 DNA=1.52pmol DNA=3.03pmol DNA 末端

1μg pBR322 DNA=0.36pmol DNA

1pmol 1000bp 长度 DNA=0.66μg DNA

1pmol pBR322 DNA=2.8μg DNA

1kb 双链 DNA 的分子量 6.6×10^5

1kb 单链 DNA 的分子量 3.3×10^5

1kb 单链 RNA 的分子量 3.4×10^5

脱氧核糖核苷酸的平均分子量 324.5

④ 蛋白质摩尔换算

100pmol 分子量 100000 蛋白质=10μg 蛋白质

100pmol 分子量 50000 蛋白质=5μg 蛋白质

100pmol 分子量 10000 蛋白质=1μg 蛋白质

氨基酸的平均分子量 126.7

⑤ 蛋白质/DNA 的换算

1kb DNA=333 个氨基酸编码容量=分子量为 3.7×10^5 蛋白质

分子量 10000 蛋白质=270bp DNA

分子量 30000 蛋白质=810bp DNA

分子量 50000 蛋白质=1.35kb DNA

分子量 100000 蛋白质=2.7kb DNA

3. 常用 DNA 分子量标准参照物

λ DNA/*Hind*Ⅲ	λ DNA/*Eco*R Ⅰ	λ DNA/*Hind*Ⅲ+*Eco*R Ⅰ	pBR322/*Hae*Ⅲ	
23123bp	21226bp	21227bp	587bp	123bp
9416bp	7421bp	5148bp	405bp	104bp
6557bp	5804bp	4973bp	504bp	89bp
4361bp	5643bp	4268bp	458bp	80bp
2322bp	4843bp	3530bp	434bp	64bp
2027bp	3530bp	2027bp	267bp	57bp
564bp		1904bp	234bp	51bp
125bp		1584bp	213bp	21bp
		1375bp	192bp	18bp
		974bp	184bp	11bp
		831bp	124bp	7bp
		564bp		
		125bp		

4. 蛋白质分子量标准

低分子量标准参照

蛋白质	分子量	蛋白质	分子量
碳酸酐酶	31000	肌球蛋白(F1)	8100
大豆胰蛋白酶抑制剂	21000	肌球蛋白(F2)	6200
马心肌球蛋白	16900	肌球蛋白(F3)	2500
溶菌酶	14400		

5. 琼脂糖凝胶浓度与线性 DNA 分辨范围

凝胶浓度/%	线性 DNA 长度/bp	凝胶浓度/%	线性 DNA 长度/bp
0.5	1000～30000	1.2	400～7000
0.7	800～12000	1.5	200～3000
1.0	500～10000	2.0	50～2000

6. 放射性核素资料

放射性核素衰变表

^{3}H		^{35}S		^{32}P		^{125}I		^{131}I	
时间/年	剩余活性/%	时间/年	剩余活性/%	时间/年	剩余活性/%	时间/年	剩余活性/%	时间/年	剩余活性/%
1	94.5	2	98.4	1	95.3	4	95.5	0.2	98.3
2	89.3	5	96.1	2	90.8	8	91.2	0.4	96.6
3	84.4	10	92.3	3	86.5	12	87.1	0.6	95.0
4	79.8	15	88.7	4	82.5	16	83.1	1.0	91.8
5	75.4	20	85.3	5	78.5	20	79.4	1.6	87.2
6	71.3	25	82.0	6	74.8	24	75.8	2.3	81.2
7	67.4	31	78.1	7	71.2	28	72.4	3.1	76.7
8	63.7	37	74.5	8	67.8	32	69.1	4.0	71.0
9	60.2	43	71.0	9	64.7	36	66.0	5.0	65.2
11	53.8	57	63.6	11	58.7	44	60.2	7.3	53.4
12	50.9	65	59.6	12	55.9	48	57.4	8.1	50.0
12.3	50.0	73	56.0	13	53.2	52	54.8		
		81	52.5	14	50.7	56	52.4		
		87.1	50.0	14.3	50.0	60	50.0		

放射性测定单位

单　位	定　义	换　算
吸收放射性剂量(拉得 rad)	100 尔格/克(电离辐射传给单位质量物质的能量)	0.87rad=1r(2.58×10^{-4}C/kg)
伦琴(r)	使 1g 空气产生 1.6×10^{12}个离子对的 X 射线或 γ 射线的照射剂量	1r=0.87rad 0.87
里(Ci)	每秒放射性核衰变 3.7×10^{10}个原子的量	1Ci=10^{3}mCi =10^{6}μCi
毫居(mCi)	千分之一居里	1mCi=10^{-3}Ci =10^{3}μCi
微居(μCi)	百万分之一居里(每分钟衰变 2.2×10^{6} 个原子的量)	1μCi=10^{-6}Ci =10^{-3}mCi
贝可勒尔(Bq)	每秒衰变一个原子的量	1Bq=3.7×10^{-10}Ci
千贝可勒尔(kBq)		1kBq=10^{3}Bq
兆贝可勒尔(MBq)		1MBq=10^{6}Bq
吉贝可勒尔(GBq)		1GBq=10^{9}Bq
每分钟衰变数(dpm)	每分钟衰变原子数	2.2×10^{6}dpm=1μCi
每分钟计数(cpm)	测量仪测出每分钟 β 粒子数	cpm=dpm×计数器效率

7. 葡聚糖凝胶的某些技术数据

分子筛类型(Sephadex)		干粒直径/μm	分子量分级范围		床体积/(ml/g)	得水值/(ml/g)	膨胀时间/h	
			球形分子	线性分子			室温	沸水
G-10		40～120	约 700	约 700	2～3	1.0±0.1	3	1
G-15		40～120	约 1500	约 1500	2.5～3.5	1.5±3.5	3	1
G-25	粗级	100～300(≈50～100 目)						
	中级	50～150(≈100～200 目)	1000～5000	100～5000	4～6	1.5±0.2	6	2
	细级	20～80(≈200～400 目)						
	超细	10～40						
G-50	粗级	100～200						
	中级	50～150	1500～30000	500～10000	9～11	5.0±0.3	6	2
	细级	20～80						
	超细	10～40						
G-75	普通	40～120	3000～70000	1000～50000	12～15	7.5±0.5	24	3
	超细	10～40	3000～70000	1000～50000	12～15	7.5±0.5	24	3
G-100	普通	40～120	4000～1500000	1000～150000	15～20	10±1.0	48	5
	超细	10～40	4000～1500000	1000～150000	15～20	10±1.0	48	5
G-150	普通	40～120	5000～400000	1000～150000	20～30	15.0±1.5	72	5
	超细	10～40	5000～400000	1000～150000	18～22	15.0±1.5	72	5
G-200	普通	40～120	5000～800000	1000～200000	30～40	30.0±2.0	72	5
	超细	10～40	5000～800000	1000～200000	20～25	30.0±2.0	72	5

8. 聚丙烯酰胺的技术数据

型　号	排阻下限	分级分离范围	膨胀后床体积/(ml/g 干胶)	膨胀所需时间(室温)/h
Bio-gel-P-2	1600	200～2000	3.8	2～4
Bio-gel-P-4	3600	500～4000	5.8	2～4
Bio-gel-P-6	4600	1000～5000	8.8	2～4
Bio-gel-P-10	10000	5000～17000	12.4	2～4
Bio-gel-P-30	30000	20000～50000	14.9	10～12
Bio-gel-P-60	60000	30000～100000	19.0	10～12
Bio-gel-P-100	100000	40000～100000	19.0	24
Bio-gel-P-150	150000	50000～150000	24.0	24
Bio-gel-P-200	200000	80000～200000	34.0	48
Bio-gel-P-300	300000	100000～400000	40.0	48

三、分子生物学研究中的网上资源

1. 一些有关核酸分析的网址和数据库

分子遗传学的引物	http://www.gdb.org/Dan/DOE/intro.html
EMBL(欧洲生物信息学研究所核酸数据库)	http://www.ebi.ac.uk/services/services.html
基因库(NCBI,美国核酸数据库)	http://www..ncbi.nlm.nih.gov/
GDB(英国医学研究会人类基因组数据库)	http://www.hgmp.mrc.ac.uk//gdb/gdbtop.html
SEQENT(英国 Daresbuy 实验室)	http://www.dl.ac.uk//gdb/gdbtop.html
甘蔗基因组信息资源	http://genome-www.standford.edu
大肠杆菌遗传贮存中心	http://cgsc.biology.yale.edu/cgsc.html
果蝇数据库(flybase)	http://morgan.harvard.edu
植物基因组数据和信息中心	http://www.nalusda.gov/answers/info-centers/pgdic
Caenorhaditis elegants 基因组计划	http://www.sanger.ac.uk/
完整的人类基因图谱	http://www.ncbi.nlm.nih.gov/SCIENCE96/

2. 一些有关蛋白质分析的网址和数据库

Swiss-Prot(EBI蛋白质序列数据库)	http://expasy.hcuge.ch/sport/sprottop.html
PIR(NBRF/MIPS蛋白质数据库)	http://www.gdb.org/Dan/proteins/pir.html
PDB(3-D蛋白质结构数据库)	http://pbl.pdb.bnl.gov/
Swiss 3-D(3-D蛋白质数据库)	http:// expasy.hcuge..ch/sw3d-top.html
LiMB(分子生物学数据库名单)	gopher://gopher.nih.gov/77/gopherlib/indices/limb
分子生物学的FTP档案	gopher://genome -gopher.stanford.edu/11/ftp
Prosite	http://expasy.hcuge.ch/sport/prosite.html
Motif服务器	http://motif.stanford.edu/emotif
蛋白质结构原理(相互作用进程)	http://www.cryst.bbk.ac.uk/pps/index.html
CHIME和RASMOL中的生物分子	http://www.umass.edu/microbio/rasmol/tutbymol.html
酶命名法数据库	http://exppasy.hcuge.ch/sport/enzyme.html
分子模型NIH中心	http://cmm.info.nih.gov/modeling/
WeblabTM分子模拟浏览器	http://www.msi.com/
ChemWeb	http://www.softshell.com/FREE/chemweb/cw-pages.html

四、细胞培养常用资料

1. 常用人工合成细胞培养基

名　称	适　用　细　胞	作　者
199	适用于人、动物非转化细胞	Morgan(1950)
BME(Basal Medium)	广泛用于人单层培养细胞	Eagle
MEM(Minimum Essential)	为当今广泛使用的培养基	Eagle
Alpha-MEM	适用于难培养的细胞	Stannerrs
McCoy-5A	适用于人与动物单层细胞和克隆细胞培养	McCoy
RPMI 1640	为McCoy-5A的改良，当前应用甚广	Moore
N16	用40%浓度，利于HeLa细胞生长	Puck
NCTC-109	胱氨酸量多	Evans
NCTC-135	为NCTC-109的改良，胱氨酸量少	Evans
CMRL-1066	为199的改良	Parker
DM-160	适合大鼠腺上皮细胞生长	Katsuta and Johnson
F12	适于克隆细胞生长	Ham
DMEM	应用甚广	Dulbecco MEM
MCDB-104;105	适于人二倍体克隆生长	Mckeehan
MB-752/1	适于小鼠细胞系	Waymouth(1959)
L-15	适于人二倍体细胞，可不用CO_2	Leibovitz(1963)
F-10	适于动物克隆细胞生长	Ham(1965)
MD705/1	为MD752/1的改良，含微量元素	Kitos(1962)
Nagle	适于哺乳动物细胞	Nagle
Higuechi	适于小鼠L细胞系	Higuchi
HIWO5	适于悬浮细胞(小鼠)	Morrison and Jenkin(1972)
F12G	适于大鼠脑细胞	Shapiro and Schier(1973)

2. 细菌、噬菌体用液

(1) 细菌液体培养基

① LB培养基 (Luria-Bertani medium)

配置每升培养基应在950ml去离子水中加入：

细菌培养用胰化蛋白胨 (bacto-tryptone)	10g
细菌培养用酵母提取物 (bacto-yeast extract)	5g
NaCl	10g

摇动容器至容器完全溶解，用5mol/L NaOH（约0.2ml）调节pH值至7.0，加入去离子水，至总体

积为 1L，高压消毒，1.034×10^5 Pa，20min。

② NZCYM 培养基

配置每升培养基应在 950ml 去离子水中加入：

NZ 胺（酪蛋白酶促水解物）	10g
NaCl	5g
细菌培养用酵母提取物（bacto-yeast extract）	5g
酪蛋白氨基酸（casamino acid）	1g
$MgSO_4$	2g

摇动容器至溶质完全溶解，用 5mol/L NaOH（约 0.2ml）调节 pH 值至 7.0，加入去离子水，至总体积为 1L，高压消毒，1.034×10^5 Pa，20min。

③ NZYM 培养基

NZM 培养基除不含酪蛋白氨基酸外，其他成分与 NZCYM 培养基相同。

④ NZM 培养基

除不含酵母提取物外，其他成分与 NZYM 培养基相同。

⑤ 高浓度肉汤

细菌培养用胰化蛋白胨（bacto-tryptone）	12g
细菌培养用酵母提取物（bacto-yeast extract）	24g
加去离子水至	800ml

搅拌，使溶质完全溶解，1.034×10^5 Pa，高压 20min，冷却至 60℃或 60℃以下，再加入 100ml 经灭菌的 0.17mol/L KH_2PO_4 0.72mol/L K_2HPO_4 溶液，该磷酸缓冲液的配法如下：

在 90ml 去离子水中溶解 2.31g KH_2PO_4，然后加入去离子水，总体积至 100ml，1.034×10^5 Pa，高压 20min。

⑥ SOB 培养基

细菌培养用胰化蛋白胨（bacto-tryptone）	12g
细菌培养用酵母提取物（bacto-yeast extract）	24g
NaCl	0.5g
加去离子水至	800ml

搅拌，使溶质完全溶解，然后加入 10ml 250mmol/L KCl 溶液，用 5mol/LNaOH 约 0.2ml 调节溶液的 pH 值至 7.4，然后加入去离子水至总体积为 1L，1.034×10^5 Pa，高压蒸汽灭菌 20min。

该溶液在使用前加入 5ml 经灭菌的 2mmol/L $MgCl_2$ 溶液；$MgCl_2$ 溶液配法如下。

在 90ml 去离子水中溶解 19g $MgCl_2$，然后加入去离子水，总体积至 100ml，1.034×10^5 Pa，高压 20min。

⑦ SOC 培养基

SOC 培养基除含有 20mmol/L 葡萄糖外，其余成分与 SOB 培养基基本相同。SOB 培养基经高压灭菌后，冷却至 60℃或 60℃以下，然后加入 20ml 经除菌的 1mol/L 葡萄糖液；1mol/L 葡萄糖溶液的配法如下。

90ml 去离子水，90ml 中加葡萄糖 18g，待完全溶解后，加入去离子水至总体积为 100ml，然后用 0.22μm 过滤膜除菌。

⑧ 2×YT 培养基

细菌培养用胰化蛋白胨（bacto-tryptone）	16g
细菌培养用酵母提取物（bacto-yeast extract）	10g
NaCl	5g
加去离子水至	800ml

搅拌，使溶质完全溶解，用 5mol/L NaOH 调节 pH 值至 7.4，然后加入去离子水至总体积为 1L，1.034×10^5 Pa，高压蒸汽灭菌 20min。

⑨ M9 培养基

Na_2HPO_4	6g
KH_2PO_4	10g
NaCl	0.5g
NH_4Cl	1g

搅拌，使溶质完全溶解，调节溶液的 pH 至 7.4，然后加入去离子水至总体积为 1L，1.034×10^5Pa，高压蒸汽灭菌 20min。待冷却后，分别加入经除菌的下列溶液：

1mol/L $MgCl_2$	2ml
20%葡萄糖	10ml
1mol/L $CaCl_2$	0.1ml

以上溶液需分别配置，葡萄糖溶液不可高压灭菌，需用 0.22μm 滤过膜除菌。

⑩ MgCA 培养基

除含 2.0g/L 酪蛋白氨基酸（casamino acid）外，其余成分与 Mg 培养基相同。

⑪ X1776 培养基

细菌培养用胰化蛋白胨（bacto-tryptone）	25g
细菌培养用酵母提取物（bacto-yeast extract）	7.5g
1mol/L Tris·HCl（pH7.6）	20ml
加去离子水至	800ml

搅拌，使溶质完全溶解，用 5mol/L NaOH 调节溶液的 pH 值至 7.4，然后加入去离子水至总体积为 1L，1.034×10^5Pa，高压蒸汽灭菌 20min。然后分别加入经除菌的下列溶液。

1mol/L $MgCl_2$	5ml
1% diaminopimelic acid	10ml
0.4%胸腺嘧啶核苷	10ml
20%葡萄糖	25ml

除 $MgCl_2$ 溶液外，其他溶液可不高压灭菌，需用 0.22μm 滤过膜除菌。

（2）含有琼脂糖的培养基

在液体培养基中加入下列试剂中的一种。

细菌培养用琼脂（bacto-agar）	15 g/L（配置平板用）
细菌培养用琼脂（bacto-agar）	7 g/L（配置顶层琼脂糖用）
琼脂糖	15 g/L（铺制平板用）
琼脂糖	7 g/L（配置顶层琼脂糖用）

1.034×10^5Pa 高压蒸汽灭菌 20min，溶液尚未冷却时，即应取出培养基，并轻轻转动以使溶解的琼脂或琼脂糖能均匀分布于整个培养基溶液中。必须小心，此时培养基溶液可能过热，旋动液体会发生暴沸，冷却至 50℃后，方可加入不耐热的物质（如抗生素）；为避免发生气泡，混匀培养基时应采取旋动的方式，然后可直接从烧瓶中倾出培养基铺制平板。90mm 直径的培养皿约需 30～50ml 培养基。培养基完全凝固后，应倒置平皿并贮存于 4℃中备用，使用前 1～2h 应取出贮存的平皿。如果平板是新鲜制备的，在 37℃温育时会“发汗”，可能导致细菌克隆或噬菌体噬斑在平板表面交互扩散而增加交叉污染的机会。为避免这一问题，可拭去平皿内部的冷凝水并把平皿倒置于 37℃中温育数小时后再使用，也可用快速猛甩一下平皿盖的方法除去冷凝水。为尽可能减少污染的机会，除去盖上的水滴时应把开盖的平皿倒持于手中。

（3）保存培养基

细菌在穿刺培养物中可保存 2 年之久；在含有甘油的培养物中可无期限保存。

① 穿刺培养物

在 2～3ml 带有螺口旋盖和橡皮垫圈的玻璃试管中，加入约 2/3 容量的熔化 LB 琼脂，旋上盖子，但并不拧紧，1.034×10^5Pa 高压 20min。冷却至室温时再拧紧盖子，置室温保存备用。保存细菌时，用一灭菌

的接种针，挑取生长良好的单菌落，把针穿过琼脂层直达瓶底数次，盖上瓶盖并拧紧，在瓶体和瓶盖上均做好标记，室温中存放于暗中。或将瓶盖拧松，37℃过夜，然后再拧紧瓶盖，于室温或4℃避光保存。

② 含甘油的培养物

在细菌培养物中加入终浓度为15%的灭菌甘油，振荡培养物使甘油分布均匀然后小量分装于带有螺口和空气密封圈的保存管内，在乙醇/干冰或液氮中冻结后，转至－70℃长期保存。复苏菌种时，用灭菌的接种针刮拭冻结的培养物表面，然后立即把黏附在接种针上的细菌划于含适当抗生素的LB琼脂平板表面。菌种管重新置于－70℃保存，琼脂平板于37℃培养过夜。

在琼脂平板上生长的细菌培养物，在琼脂平板表面刮下细菌放入装有2ml LB的无菌试管内，再加入等量的含有30%灭菌甘油的LB培养基，振荡混合物使甘油完全分布均匀后，分装于带有螺口盖和空气密封圈的无菌试管中，按上述方法冰冻保存。

(4) 抗生素

抗生素	贮存液①		工作浓度	
	浓 度	保存条件	严紧型质粒	松弛型质粒
氨苄青霉素	50mg/ml(溶于水)	－20℃	20μg/ml	60μg/ml
羧苄青霉素	50mg/ml(溶于水)	－20℃	20μg/ml	60μg/ml
氯霉素	34mg/ml(溶于乙醇)	－20℃	25μg/ml	170μg/ml
卡那霉素	10mg/ml(溶于水)	－20℃	10μg/ml	50μg/ml
链霉素	10mg/ml(溶于水)	－20℃	10μg/ml	50μg/ml
四环素②	5mg/ml(溶于水)	－20℃	10μg/ml	50μg/ml

① 以水为溶液剂的抗生素贮存液应通过0.22μm滤器滤过除菌，以乙醇为溶剂的抗生素溶液无需除菌处理，所有抗生素溶液均应放于不透光的容器中保存。

② 镁离子是四环素的拮抗剂，四环素抗性菌的筛选应使用不含镁盐的培养基（如LB培养基）。

3. 细胞培养中出现的常见问题、原因及解决方法

问　题	原　因	解决办法
培养液pH迅速变化	CO_2浓度不正确；液体中碳酸氢盐浓度不足；细菌、酵母菌、霉菌污染	增加或降低CO_2浓度(5%～10%)；碳酸氢盐浓度2.0～3.7g/L，HEPES浓度10～25mmol/L，丢弃污染的培养液
液体有沉淀，pH不变	瓶底残留盐类物质发生沉淀	重新过滤培养液，更新容器
	培养液经过冻融	加温可溶解，否则丢弃
液体有沉淀，pH变低	微生物污染	丢弃培养物
细胞不贴壁	胰酶消化过度；支原体污染	缩短胰酶消化时间；丢弃污染的细胞并清洁超净台和培养箱；增加贴壁因子
细胞生长缓慢	培养基或血清型号不对；生长因子、谷氨酰胺不足或失效；培养液保存不当或失效；低水平霉菌污染或支原体污染；细胞老化	核对培养基；应逐步换新血清；更换新培养液；重新加入谷氨酰胺等；血清应保存于－20℃；配好的培养基应贮存于2～8℃，并应尽可能避光保存；不使用过期的培养液 丢弃污染的细胞并清洁超净台及培养箱更换新的细胞
细胞死亡	培养环境缺少CO_2；温度不恒定；抗生素浓度过高；细胞冻存或复苏时损伤；培养液中渗透压不准确；培养液中含有毒代谢物	核对培养基浓度，CO_2浓度及抗生素；更换新培养液；哺乳动物细胞培养液渗透压为260～350mmol/kg；HEPES等可影响渗透压
细胞成团	钙、镁离子引起；支原体污染；蛋白酶的过度消化导致DNA释放及细胞裂解	无钙、镁离子的Hank's液冲洗细胞，用吸管将悬浮培养细胞轻轻吹打成单个；丢弃污染的细胞；用DNase Ⅰ处理细胞

参 考 文 献

1 陈竺，张伯勤，方福德. 基因组科学与人类疾病. 北京：科学出版社，2002
2 梁国栋. 最新分子生物学实验技术. 北京：科学出版社，2001
3 贺林等. 解码生命. 北京：科学出版社，2000
4 Sambrook J，Fritsch E F，Maniati T. Molecnlar Cloning. Cold Spring Harbor Laboratory Bress. USA. 1992
5 卢圣栋. 现代分子生物学实验技术. 北京：中国协和医科大学出版社，1999
6 林万明. PCR 技术操作和应用指南. 北京：人民军医出版社，1993
7 李建武. 生物化学实验原理与方法. 北京：北京大学出版社，1994
8 孙志贤. 现代生物化学理论与研究技术（研究生教材）. 北京：军事医学科学出版社，1995
9 夏其昌. 蛋白质化学研究技术与进展. 北京：科学出版社，1999
10 周爱儒. 医学生物化学. 北京：北京医科大学出版社，1997
11 杰里米·里夫金. 生物技术世纪. 上海：上海科技教育出版社，2000
12 王联结. 生物化学与分子生物学. 北京：科学出版社，2004
13 李元宗. 生化分析. 北京：高等教育出版社，2003
14 陈均辉. 生物化学实验. 北京：科学出版社，2004
15 赵永芳. 生物化学技术原理与应用. 北京：科学出版社，2004
16 Mark Schena. DNA Microarrays：a practical references. New York：Oxford University Press Inc，2001
17 马立人，蒋中华. 生物芯片. 北京：化学工业出版社，2000
18 温进坤，韩梅. 医学分子生物学理论与研究技术. 第 2 版. 北京：科学出版社，2002.
19 Latchman 等 D S. 转录因子实用技术. 第 2 版. 刘进元，赵广荣等译. 北京：清华大学出版社，2004
20 周爱儒. 生物化学. 第 5 版. 北京：人民卫生出版社，2001
21 吴乃虎. 基因工程原理. 第 2 版. 北京：科学出版社，2001
22 冯作化. 医学分子生物学. 北京：人民卫生出版社，2001
23 张玉静. 分子遗传学. 北京：科学出版社，2000
24 凌诒萍. 细胞生物学. 北京：人民卫生出版社，2001
25 楼士林等. 基因工程. 北京：科学出版社，2002
26 谷志远. 现代医学分子生物学. 北京：人民军医出版社，1998
27. 静国忠. 基因工程及其分子生物学基础. 北京：北京大学出版社，1999
28 孙树汉. 基因工程原理与方法. 北京：人民军医出版社，2001
29 孙含笑等. 转基因技术理论与应用. 郑州：河南医科大学出版社，2000
30 张学敏等. 靶向新基因的分子克隆策略——理论与方法. 北京：军事医学科学院出版社，1999
31 沈倍奋. 分子文库. 北京：科学出版社，2001
32 毛忠贵. 生物工业下游技术. 北京：中国轻工业出版社，1999
33 冯伯森. 动物细胞工程原理与实践. 北京：科学出版社
34 方福德等. 分子生物学前沿技术. 北京：北京医科大学中国协和医科大学联合出版社，1998
35 张惠展. 基因工程概论. 上海：华东理工大学出版社，1999
36 刘志国. 新编生物化学. 北京：中国轻工业出版社，2003
37 李育阳. 基因表达技术. 北京：科学出版社，2001
38 特怀曼 RM. 高级分子生物学要义. 北京：科学出版社，2000
39 刘谦. 生物安全. 北京：科学出版社，2001
40 Murray R K，Granner D K，et al. Harper's Biochemistry 25th ed.（英文影印版）. 北京：科学出版社
41 Alberts，et al. Molecular Biology of the cell. Garland pub Inc，1994
42 Meisenberg G and Simmons W H. Principles of Medical Biochemistry St. Louis：Mosby Inc，1998
43 Lodish H，et al. Molecular Cell Biology. Freeman W H and Company，2000
44 Brent R. Genomic Biology. Cell，2000，100：169～183
45 Lander E S，Linten L M，Birren B，et al. International Human Genome Sequencing Consortium. Nature，2001，409

(6822): 860～921

46 Venter J C, Adams E W, et al. The Sequence of the human genome. Science, 2001, 291 (5507): 1304～1351

47 Chen Y J, et al. Genes Chromosomes Cancer, 1999, 25 (2): 169～175.

48 Wessendorf S, et al. Lab Invest, 2002, 82 (1): 47～60

49 Clark J, et al. Genes Chromosomes Cancer, 2002, 34 (1): 104～114

50 Veltman J A, et al. Am J Hum Genet, 2002, 70 (5): 1269～1276

51 Hui A B, et al. Int J Onco, 2002, 20 (3): 467～473

52 Lash A E, Tolstoshev C M, Wagner L, et al. Genome Res, 2000, 10 (7): 1051～1060

53 Larsson M, Stahl S, Uhlen M, et al. Genomics, 2000, 63 (3): 341～353

54 Riggins G J, Strausberg R L. Hum Mol Genet, 2001, 10 (7): 663～667

55 Lee S, Zhou G, Clark T, et al. Proc Natl Acad Sci, 2001, 98 (6): 3340～3345

56 Reprogen Inc, et al. Genomics, 2000, 63 (1): 13～24

57 Frohman M A. Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE [J]. Method Enzymol, 1993, 218: 340～356

58 生秀杰. 基因打靶的策略及其发展. 国外医学·遗传学分册. 2001, 24 (1): 8～10

59 李子银等. 遗传, 2000, 22 (1): 57～60

60 Springer P S, et al. Plant Cell, 2000, 12 (7): 1007～1020

61 Mitchell K J, et al. Nat Genet, 2001, 28 (3): 198～200

62 Zhennan Lai, et al. PNAS, 2002, 99 (6): 3651～3656

63 Brantl S, et al. Journal of Bacteriology, 2002, 184 (10): 2740～2747

64 Lebedeva I, et al. A nnu Rev Pharmacol Toxicol, 2001, 41: 403～419

65 Childs J L, et al. Proc Natl Acad Sci, 2002, 99 (17): 11091～11096

66 Brantl S. Antisense-RNA regulation and RNA interference. Biochem Biophys Acta, 2002, 1575 (123): 15～25

67 Zamore P D. Ancient pathways programmed by small RNAs. Science, 2002, 296 (5571): 1265～1269

68 Hannon G J. RNA interference. Nature, 2002, 418 (6894): 244～251

69 Doench J G, Petersen C P, Sharp P A. siRNAs can function as miRNAs. Genes Dev, 2003, 17 (4): 438～442

70 Plasterk R H. RNA silencing: the genome's immune system. Science, 2002, 296 (5571): 1263～1265

71 Sayda M E, Jens H, Winfried L, et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in all cultured mammalian cells [J]. Nature, 2001, 411: 494～498

72 Michael T, McManus, Phillip A S. Gene silencing in mammals by small interfering RNAs. Nature Reviews Genetics, 2002, 3: 737～747

73 David B. RNA silencing: Diced defence. Nature, 2001, 409: 295～298

74 陆春叶. 人类单核苷酸多态性及其应用. 国外医学·遗传学分册, 2001, 24 (1): 21～24

75 陈炜, 张戈, 张思仲. 基于生物信息学的 SNP 候选位点搜寻方法. 遗传, 2001 (2): 153～156

76 Huttenhofer A, et al. Rnomics: an experimental approach that identifies 201 candidates for novel, small, non-message RNAs in mouse. The EMBO J, 20 (11): 2943～2953

77 Cheng J, Waters L C, Fortina P, et al. Degenerate oligonucleotide primed-polymerase chain reaction and capillary electrophoretic analysis of human DNA on microchip-based devices. Anal Biochem, 1998, 257 (2): 101～106

78 Cheng J, Shoffner M A, Hvichia G E, et al. Chip PCR. Ⅱ. Investigation of different PCR amplification systems in microbabricated silicon-glass chips. Nucleic Acids Res, 1996, 24 (2): 380～385

79 Hacia J G. Resequencing and mutational analysis using oligonucleotide microarrays. Nat Genet, 1999, 21 (1 Suppl): 42～47

80 Lueking A, Horn M, Eickhoff H, et al. Protein microarrays for gene expression and antibody screening. Ansl Biochem, 1999, 270 (1): 103～111

81 Bains W. Selection of oligonucleotide probes and experimental conditions for multiplex hybridization experiments. Genet Anal Tech Appl, 1994, 11 (3): 49～62

82 Honore B, Madsen P. The tetramethylammonium chloride (TMAC) method for screening cDNA libraries with highly

degenerate oligonucleotide probes obtained by reverse translation of amino acid sequences. Methods Mol Biol, 1997; 69: 139～146

83 Honore B, Madsen P, Leffers H. The tetramethylammonium chloride method for screening of cDNA libraries using highly degenerate oligonucleotides obtained by backtranslation of amino-acid sequences. J Biochem Biophys Methods, 1993, 27 (1): 39～48

84 Oh S H, Fleischhauer K, Yang S Y. Isoelectric focusing subtypes of HLA-A can be defined by oligonucleotide typing. Tissue Antigens, 1993, 41 (3): 135～142

85 Nelson R W, Nedelkov D, Tubbs K A. Biosensor chip mass spectrometry: A chip-based proteomics approach. Eletrophoresis, 2000, 21, 1155～1163

86 Sun Z, Fu X, Zhang L, et al. A protein chip system for parallel analysis of multi-tumor markers and its application in cancer detection. Anticancer Res, 2004, 24 (2C): 1159～1165

87 Melton L. Protein array: proteomics in multiplex. Nature, 2004, 429 (6987): 101～107

88 Hodgetts A, Bosse J T, Kroll J S, et al. Analysis of differential protein expression in Actinobacillus pleuropneumoniae by Surface Enhanced Laser Desorption Ionisation——ProteinChip (SELDI) technology. Vet Microbiol, 2004, 99 (3～4): 215～225

89 Vlahou A, Schellhammer P E, Wright G L Jr. Application of a novel protein chip mass spectrometry technology for the identification of bladder cancer-associated biomarkers. Adv Exp Med Biol, 2003, 539 (Pt A): 47～60

90 Balachandra K, Laisupasin P, Dhepakson P, et al. Preliminary clinical evaluation of a protein chip for tumor marker serodiagnosis of various cancers. Asian Pac J Allergy Immunol, 2003, 21 (3): 171～178

91 Gershon D. Proteomics technologies: probing the proteome. Nature. 2003, 424 (6948): 581～587.

92 Valle R P, Jendoubi M. Antibody-based technologies for target discovery. Curr Opin Drug Discov Devel, 2003, 6 (2): 197～203

93 Kusnezow W, Hoheisel J D. Antibody microarrays: promises and problems. Biotechniques. 2002, Suppl: 14～23

94 James P. Chips for proteomics: a new tool or just hype? Biotechniques. 2002, Suppl: 4～10, 12～13

95 Agrawal N, et al. RNA interference: biology, mechanism, and applications. Microbiol Mol Biol Rev, 2003, 67 (4): 657～685

96 Narry Kim. V. RNA interference in function genomics and medicine. J Korean Med Sci, 2003, 18: 309～318

97 Vaucheret, et al. Post-transcriptional gene silencing in plants. J Cell Science, 2001, 114 (17): 3083～3091

98 Fire, et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature, 1998, 391 (6669) 806～811

99 Elbashir S M, et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature, 2001, 411: 494～498

100 Elbahir S M, et al. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. EMBO J, 2001, 20: 6877～6888

101 Elbashir S M, et al. Analysis of gene function in somatic mammalian cells using small interfering RNAs. Methods, 2002, 26: 199～213

102 Reynolds A, Leake D, Boese Q, Scaringe S, Marshall W S, Khvorova A. Rational siRNA design for RNA interference. Nat Biotechnol. 2004, 22 (3): 326～330

103 Thomas Tuschl. Targeting genes expressed in mammalian cells using siRNAs. Nature Methods, 2004, Vol X (X): 13～17

104 Patel P S, Yang S Li A, Varney M L Singh R K. Capsaicin up-regulates vascular endothelial cell growth factor expression by induction of hypoxia inducing factor-1. in human malignant melanoma cells. Journal of Cancer Research and Clinical Oncology, 2002, 128: 461～468

105 Pfeifer G P, Tanguay R L, Steigerwald S D, Riggs A D. In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. Genes Dev, 1990, 4 (8): 1277～1287

106 Shashwati Basak and Nagaraja V. A versatile in vivo footprinting technique using 1,10-phenanthroline-copper complex

to study important cellular processes. Nucleic Acids Research, 2001, 29 (21): 105

107 Chang-Jang Kang, et al. Dynamic chromatic remodellinf in vicity of J chain gene for the regulation of two stages-specific genes during B cell differentiation. Mol Cells, 2001, 10 (1) 32～37

108 Michael Brenowitz. et al. Probing the structural dynamics of nuxleic acid by quantitative time-resolved and equilibrium hydroxyl radical "footprinting". Current Opinion in Structural Biology, 2002, 12: 648～653

109 Nolo R, Abbott L A, Bellen H J. Senseless, a *Zn finger transcription factor*, is necessary and sufficient for sensory organ development in Drosophila. Cell, 2000, 102 (3): 349～362